IJCB
INTERNATIONAL JOURNAL OF CENTRAL BANKING
Volume 19, Number 3
August 2023

Volume 19, Number 3 August 2023

IJCB

INTERNATIONAL JOURNAL OF CENTRAL BANKING

IJCB International Journal of Central Banking

Volume 19, Number 3 August 2023

The contents of this journal, together with additional materials provided by article authors, are available without charge at ***www.ijcb.org***.

The *International Journal of Central Banking* is published bimonthly (ISSN: 1815-4654). Online access to the publication is available free of charge at **www.ijcb.org**.

Requests for permission to reprint material from this journal should be addressed to:

International Journal of Central Banking
Printing & Fulfillment K1-120
Federal Reserve Board
Washington, DC 20551

Phone: 202-452-3425
Fax: 202-728-5886
E-mail: editor@ijcb.org

ISSN: 1815-4654

Modeling the Asymmetric Effects of an Oil Price Shock*

Lance J. Bachmeier[a] and Benjamin D. Keen[b]
[a]Kansas State University
[b]University of Oklahoma

This paper documents that an oil price increase generates a larger decline in output when the oil price hits a near-term high. We develop a New Keynesian model with energy and a downward nominal wage rigidity that generates asymmetric responses of the macroeconomy to energy price shocks. Specifically, a large energy price increase pushes down the real wage enough that the downward nominal wage constraint binds for several periods, which causes firms to reduce their output further. Since that mechanism is unimportant when energy prices fall, the downward nominal wage constraint causes output to react asymmetrically to oil price shocks.

JEL Codes: E32, Q43.

1. Introduction

Many empirical studies have documented that oil price shocks have a negative effect on output.[1] One key finding in the literature is

*We would like to thank Nathan Balke, James MacGee, Nida Cakir Melek, Michael Plante, Martin Seneca, Martin Stuermer, and Robert Vigfusson for many helpful discussions and comments. The paper also benefited from comments by seminar and conference participants at the Bank of Finland, Colgate University, Federal Reserve Bank of Dallas, Federal Reserve Bank of Kansas City, University of Nevada-Reno, Wichita State University, 2017 Southern Economic Association Meetings, 2018 International Conference on Computing in Economics and Finance, 2019 Society for Nonlinear Dynamics and Econometrics Symposium, and 2019 International Journal of Central Banking Research Conference.
Author Contact: Lance J. Bachmeier, Department of Economics, Kansas State University, Waters Hall 321, Manhattan, KS 66506. Tel: (785) 532-4578. E-mail: LanceB@ksu.edu. Benjamin D. Keen, Department of Economics, University of Oklahoma, 308 Cate Center Drive, 437 Cate Center One, Norman, OK 73019. Tel: (405) 325-5900. E-mail: ben.keen@ou.edu.

[1]For reviews of the early literature, see Hamilton (2008) and Kilian (2008). A few recent contributions include Kilian (2009), Hamilton (2011), Kilian and Vega

that models with asymmetry or another form of non-linearity fit the data better and provide superior forecasts compared with linear VAR models. Two of the most popular specifications are Hamilton's (1996, 2003, 2011) "net oil price increase" model and Kilian and Vigfusson's (2013) "net oil price change" model. The net oil price increase model predicts that a rise in oil prices generates a larger decline in output when the price of oil hits a near-term high relative to its recent history.[2] In contrast, the net oil price change model claims that a change in oil prices generates a larger shift in output when the price of oil hits either a near-term high or a near-term low relative to its recent history. In previous theoretical research, oil price shocks were unable to generate an output response consistent with either the net oil price increase model or the net oil price change model. This paper develops a New Keynesian model with energy and a downward nominal wage rigidity. A particularly interesting result of our model is that output responds asymmetrically to an energy price shock. Specifically, a large energy price increase has a greater effect on output than a large energy price decrease of the same magnitude.

A few theoretical models have been used to motivate the asymmetric responses of output to oil price changes. Bernanke (1983) suggests that agents reduce their irreversible investment whenever an exogenous shock, such as a large oil price change, increases economic uncertainty. The asymmetry in that framework, however, depends on the uncertainty generated by the price change and not the direction of the price change. Hamilton (1988) argues that capital and labor cannot costlessly move from the sectors that experience a decline in demand to the sectors that experience an increase in demand. That lack of mobility means output will definitely fall after an oil price increase, and it may even fall after an oil price decrease (Hamilton 2003). Although Mork (1989) finds some

(2011), and Aastveit (2014). Papers that have worked with theoretical models of the effects of oil price shocks similar to this paper include Bodenstein, Erceg, and Guerrieri (2008), Dhawan and Jeske (2008), Blanchard and Gali (2010), Dhawan, Jeske, and Silos (2010), Blanchard and Riggi (2013), Bodenstein, Guerrieri, and Gust (2013), Plante (2014), Gavin, Keen, and Kydland (2015), Balke and Brown (2018), and Drechsel and Tenreyro (2018).

[2]Hamilton (2003) finds that comparing the current oil price with its values over the previous three years fits the data best.

empirical support for Hamilton's (1988) costly reallocation of resources argument, Herrera, Lagalo, and Wada (2011) and Kilian and Vigfusson (2011) find Hamilton's (1988) theoretical explanation inconsistent with asymmetries observed in the data. Wei (2003) uses a general equilibrium model with putty-clay investment to show higher oil prices amplify the decline in output by making some capital obsolete. The putty-clay model, however, does not allow for the substitutability of the factors of production once capital is installed, which means Wei's specification has some of the characteristics of Hamilton's (1988) costly reallocation of resources model.

We begin by documenting the asymmetric effects of an oil price shock using a two-regime model in which the economy is in the "high oil price regime" when the price of oil hits a near-term high but is otherwise in the "normal oil price regime."[3] Our results show that an oil price increase reduces output more in the high oil price regime than in the normal oil price regime. An oil price increase in the high oil price regime also affects the labor market by generating higher nominal wages and lower hours worked. In the goods market, consumption, business fixed investment, and non-residential investment all decline more rapidly following an oil price increase in the high oil price regime rather than in the normal oil price regime.

This paper develops a New Keynesian model with downward rigid nominal wages in which an energy price increase generates asymmetric effects in the goods and labor markets consistent with our empirical observations.[4] In our model, energy is both an input in the production function and a consumption good, where the constraint preventing nominal wages from falling is needed to generate asymmetric effects after a large energy price shock. Specifically, downward rigid nominal wages enhance the decline in output after a large energy price increase by preventing the nominal wage from falling. The increase in energy prices drives up production costs, which causes firms to reduce their labor demand. Higher energy

[3]Our choice of a non-linear and asymmetric specification for the empirical model is motivated by the theoretical model developed in Section 3.

[4]The empirical literature on downward rigid nominal wages includes Gottschalk (2005), Barattieri, Basu, and Gottschalk (2014), and Hazell and Taska (2018), while the theoretical literature includes Kim and Ruge-Murcia (2009, 2011), Benigno and Ricci (2011), Abbritti and Fahr (2013), Abo-Zaid (2013), Schmitt-Grohe and Uribe (2016), and Baqaee (2020).

costs also decrease households' demand for energy, consumption, and investment, but they increase households' supply of labor. The reduction in labor demand combined with the increase in labor supply puts downward pressure on real and nominal wages. When the pressure is strong enough, the nominal wage hits its downward constraint and is unable to decline any further. Firms respond by reducing their labor demand and output by more than they would have in a flexible wage economy. As a result, a New Keynesian model with downward rigid nominal wages generates asymmetric effects after a large energy price increase. It is important to understand that the presence of the downward nominal wage rigidity is not, by itself, sufficient to generate asymmetric responses to energy price shocks. An energy price shock will only produce asymmetric responses when the downward nominal wage rigidity binds. If the energy price is at or below its steady state, a small increase in energy prices will not produce asymmetric effects because the decline in the real wage is not large enough to cause the downward nominal wage rigidity to bind.[5]

The paper proceeds as follows. Section 2 describes the data, methodology, and impulse response functions for key variables after an oil price shock in both the high and normal oil price regimes. Section 3 presents our theoretical model. Section 4 discusses the calibration and the solution technique for our theoretical model. Section 5 displays the model's impulse response functions to an energy price increase and decrease, illustrates the decision rules associated with various sizes of energy price shocks, examines the robustness of our results to alternative calibrations of key parameters, compares the effects of an energy price shock in the 1970s with that in the 2000s, and discusses the differences in the effects of an energy price shock caused by foreign demand and supply shocks. Section 6 concludes.

2. Stylized Facts

This section documents the observed effects of oil price shocks on key economic variables that any plausible theoretical model of the

[5]This statement assumes that the steady-state headline inflation rate, π^*, is greater than the degree of the downward nominal wage rigidity, γ, (see Equation (6)). If, on the other hand, γ equals π^*, then the model will generate asymmetric effects even following a small energy price increase.

transmission of oil price shocks to the macroeconomy should be able to replicate. Specifically, we show that large oil price increases that push oil prices to near-term highs have a much greater relative impact on output than other oil price shocks. The logic behind this asymmetry is that firms and households often adjust their behaviors only when the price of oil hits levels it has not reached in recent years. Those same firms and households, however, are prepared to manage the expected day-to-day fluctuations that characterize the price of oil, so daily changes have minimal economic impact.

Hamilton's (1996, 2003, 2011) net oil price increase model argues that the difference between the current oil price, oil_t, and the highest oil price in the last 12 quarters, $\widetilde{oil}_t = \max(oil_{t-1}, \ldots, oil_{t-12})$, has a larger impact on economic activity when the current price of oil rises above its three-year high, $oil_t > \widetilde{oil}_t$. Such a specification is a transformation that eliminates most of the variation of the oil price series. Applying the logic of Blanchard and Gali (2010), a downward rigid nominal wage constraint causes non-linearity similar to the net oil price increase model, but it does not transform the oil price data. The Blanchard and Gali model argues that firms respond to a rise in the oil price by reducing output and wages. If nominal wages are bound by a constraint that prevents them from falling, then firms are forced to cut output further.

Our empirical model defines the economy to be in the *high oil price regime* if $oil_t > \widetilde{oil}_t$ and in the *normal oil price regime* if $oil_t \leq \widetilde{oil}_t$. We set the dummy variable $high_t$ equal to 1 in the high oil price regime and equal to 0 in the normal oil price regime. Since the value of $high_t$ depends on the size of oil_t relative to $\widetilde{oil}_t$, $high_t$ is effectively a threshold dummy variable that introduces non-linearity into the model. The dummy variable interacts with the price of oil, $H_t = high_t \Delta oil_t$, to measure the additional impact of an oil price increase in the high oil price regime compared with the same-sized increase in the normal oil price regime. Thus, $H_t > 0$ in the high oil price regime, whereas $H_t = 0$ in the normal oil price regime.[6]

[6]Our model is not the same as Hamilton's (1996, 2003, 2011) net oil price increase model. Although we use Hamilton's measure to classify the oil price regime as a high or normal regime, we do not apply his non-linear transformation to the oil price series. The difference is in the calculation of the oil shock in the high oil price regime. Hamilton's oil shock is the percentage change of the

Figure 1 plots separately the oil price changes in the high oil price regime, $high_t = 1$, and the normal oil price regime, $high_t = 0$, from 1972:Q2 through 2017:Q4.[7] The economy is in the high oil price regime for 30 percent of the time and in the normal oil price regime for the remaining periods. Our focus is on the large oil price increases in the high oil price regime because those shocks can have a substantial effect on output. Figure 1 reveals that oil prices rose swiftly above their recent highs in 1973–74, 1979–80, 1981, 1990, and the early 2000s, with each period being followed by a recession. The first four oil price increases are usually attributed to foreign oil supply disruptions, while the 2002–08 oil price spike is often credited to higher oil demand from China and India. Most economists believe those negative oil supply shocks were either the primary reason for or a key contributing factor of the 1974–75, 1980, 1981–82, and 1991 recessions, whereas the financial crisis was the primary cause of the 2008 recession, as opposed to the large rise in oil demand. Given that those large oil price increases significantly affected output, our paper estimates the economic effects of large oil price increases and then builds a theoretical model that generates responses to energy price shocks consistent with that behavior.

2.1 Methodology

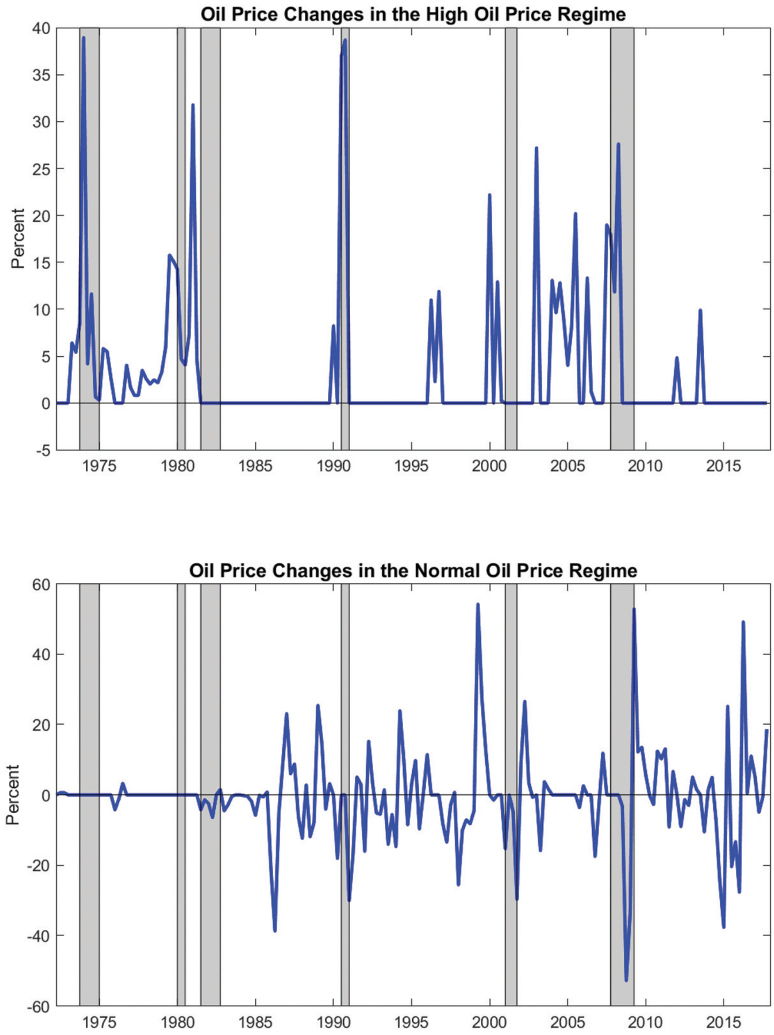
The threshold dummy variable $high_t$ introduces a non-linearity into our model. We estimate the model using the method of local projections introduced by Jorda (2005). Our empirical analysis then compares the impulse response functions from a 9.5 percent oil price increase in the high oil price regime with the impulse response functions from a 9.5 percent oil price increase in the normal oil price regime.[8] This model differs somewhat from the literature on state-dependent effects of fiscal and monetary policy (e.g., Ramey and Zubairy 2018) in that the variable being shocked also determines the state and therefore changes in regime.

current price of oil over the recent maximum price of oil. That comparison is not consistent with the New Keynesian model we present later, so we use the percentage change in the price of oil without doing a transformation.

[7]The oil price data is defined in Section 2.2.

[8]9.5 percent is the average quarterly increase in the price of oil in the high oil price regime.

Figure 1. Oil Price Changes in the High and Normal Oil Price Regimes



Note: Shaded areas represent recessions. An oil price change of zero indicates either the economy is not in that oil price regime or the oil price was unchanged.

Kilian (2009) and Kilian and Vega (2011) present strong evidence that oil prices are predetermined with respect to current economic conditions.[9] Therefore, we estimate the impact of an oil price change in period $t-s$ on key macroeconomic variables in period t,

$$\Delta x_t = \theta_s + \sum_{i=0}^{p} \delta_{s,i} \Delta oil_{t-s-i} + \sum_{i=0}^{p} \phi_{s,i} H_{t-s-i} + \varepsilon_{s,t}, \qquad (1)$$

where Δx_t is the percentage change in the macroeconomic variable, Δoil_t is the percentage change in the price of oil, and $\varepsilon_{s,t}$ is the error term.[10] A separate estimate of (1) is calculated for each forecast horizon, s, where $s = 0, \ldots, 8$. At each horizon s, the estimated values $\widehat{\delta}_{s,0}$ and $\widehat{\phi}_{s,0}$ are multiplied by $\Delta oil^H_{t-s} = 0.095$ and $H^H_{t-s} = 0.095$ in the high oil price regime and $\Delta oil^N_{t-s} = 0.095$ and $H^N_{t-s} = 0$ in the normal oil price regime to generate the s-period impulse response for $\Delta\widehat{x}^H_{t,s}$ in the high oil price regime and $\Delta\widehat{x}^N_{t,s}$ in the normal oil price regime, respectively. We then calculate the s-period cumulative impulse response functions, $CR_{t+s} = \sum_{j=0}^{s} \Delta\widehat{x}_{t,j}$, to present all of the variables, except the inflation rate, in level form.[11]

We are interested in determining whether oil price shocks affect key economic variables symmetrically, $\widehat{\phi}_{s,0} = 0$, or asymmetrically, $\widehat{\phi}_{s,0} \neq 0$. In the case of symmetry, the cumulative impulse response functions from an oil price increase in the high oil price regime, CR^H_{t+s}, are equal to the cumulative impulse response functions from an oil price increase in the normal oil price regime, CR^N_{t+s}. Oil price shocks, however, have asymmetric effects in the high oil price regime

[9] We also assume that the state of oil prices is predetermined with respect to economic conditions.

[10] We use the current value and three lags of the oil price as controls, so oil price data from the last full year affect the macroeconomic variables. Since there is no clear procedure to choose the number of lags when using Jorda's (2005) method of local projections, we select enough lags to capture the predictable effects of oil price shocks, while avoiding inefficient estimation by including too many lags. Proper lag selection requires a model of Δoil_t to obtain the best possible identification of the oil shock. An advantage of the local projections method is that it does not require a fully specified structural VAR model.

[11] The cumulative impulse response functions convert the period-by-period percentage changes to the percentage deviations of that data from their long-run levels.

when there is a significant difference between the cumulative impulse response functions CR^H_{t+s} and CR^N_{t+s}. That is, the cumulative difference function, $CD_{t+s} = CR^H_{t+s} - CR^N_{t+s}$, is significantly different than zero. To determine if oil price shocks have asymmetric effects, we construct the 95 percent confidence intervals for the cumulative difference functions by using the fact that H_{t-s} is the only regressor that is different in the two regimes (i.e., $H^H_{t-s} = 0.095$ and $H^N_{t-s} = 0$). Specifically, we multiply the heteroskedasticity and autocorrelation consistent (HAC) standard errors on $\widehat{\phi}_{s,0}$ by H^H_{t-s} to get the standard errors for the differences in the impulse response functions, $\Delta x^H_{t+s} - \Delta x^N_{t+s}$. The standard errors for $\Delta x^H_{t+s} - \Delta x^N_{t+s}$ are combined using the delta method to generate the standard errors for the cumulative difference functions and their resulting 95 percent confidence intervals.[12]

2.2 Data

Table 1 displays the data and their mnemonics. Each of the data series is transformed into its quarterly percentage change. The crude oil price data was obtained from the Bureau of Labor Statistics. All of the other data were downloaded from the Federal Reserve Bank of St. Louis' FRED database. Impulse response functions were computed using data over the period of 1972:Q1–2018:Q1. That sample period was chosen to avoid the inflated effects of oil price shocks when using data prior to the early 1970s.[13]

2.3 Empirical Impulse Response Functions

Figures 2 through 5 present the cumulative impulse response functions and the cumulative difference functions for output, wages, hours worked, investment, consumption, and inflation following a

[12]The covariances of coefficients across equations are ignored due to the difficulty of correcting for serial correlation when using multiple values of s in the same system. We did, however, apply a pairs bootstrap to compute the cross-equation covariances. The change has no meaningful impact on the results because the off-diagonal elements of the covariance matrix are much smaller in magnitude than the variance terms.

[13]See Herrera, Lagalo, and Wada (2011).

Table 1. The Data (Mnemonics)

Producer Price Index by Commodity for Fuels and Related Products and Power: Crude Petroleum (WPU0561)
Real Gross Domestic Product (GDPC1)
Industrial Production (INDPRO)
Industrial Production: Durable Manufacturing (IPDMAN)
Hourly Earnings: Private Sector for the United States (LCEAPR01USQ189S)
Average Hourly Earnings of Production and Non-supervisory Employees: Manufacturing (CES3000000008)
Non-farm Business Sector: Average Weekly Hours (PRS85006022)
Average Weekly Hours of Production and Non-supervisory Employees: Manufacturing (AWHMAN)
Real Gross Private Domestic Investment: Fixed Investment (A007RL1Q225SBEA)
Real Gross Private Domestic Investment: Fixed Investment: Residential (A011RL1Q225SBEA)
Real Gross Private Domestic Investment: Fixed Investment: Non-residential (A008RL1Q225SBEA)
Real Private Fixed Investment: Nonresidential: Structures: Mining Exploration, Shafts, and Wells (E318RL1Q225SBEA)
Real Personal Consumption Expenditures (DPCERL1Q225SBEA)
Real Personal Consumption Expenditures Excluding Food and Energy (DPCCRL1Q225SBEA)
Personal Consumption Expenditures: Chain-type Price Index Less Food and Energy (JCXFE)
Personal Consumption Expenditures: Chain-type Price Index (PCECTPI)

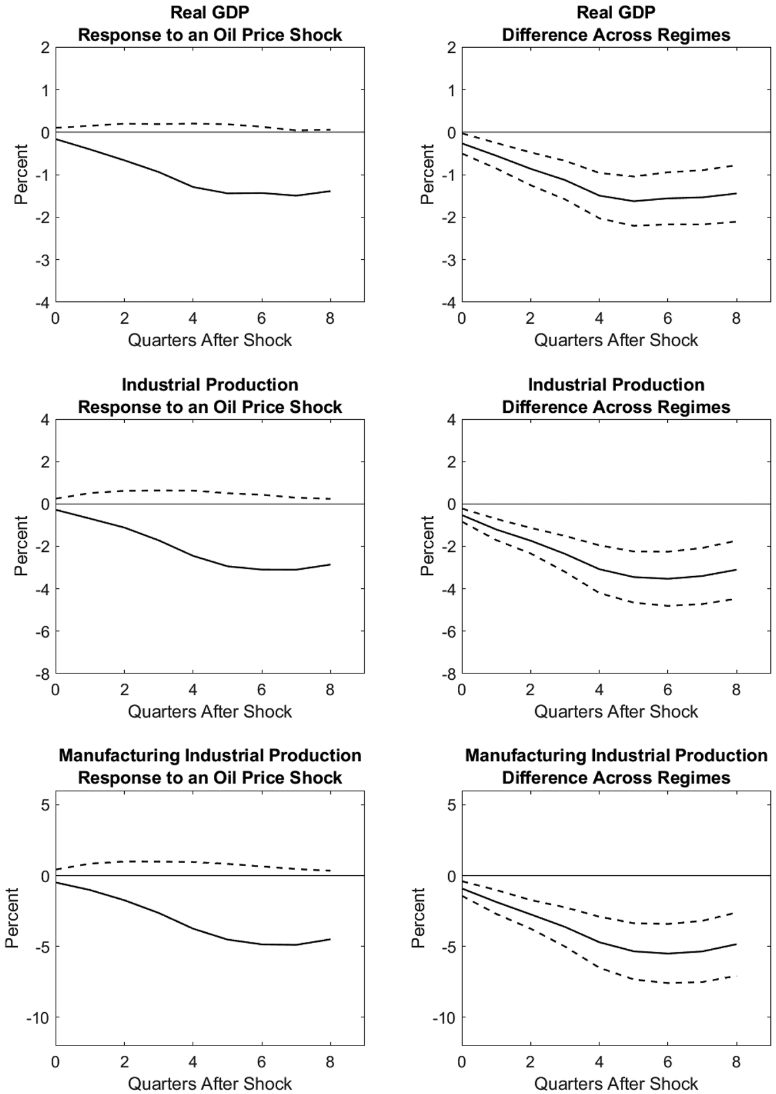
9.5 percent oil price increase.[14] In the left-hand column of each figure, the solid lines represent the cumulative impulse responses from the oil price shock in the high oil price regime, and the dashed lines display the cumulative responses following the same shock in the normal oil price regime. In the right-hand column of each figure, the solid line represents the cumulative difference functions, and the dashed lines show their 95 percent confidence intervals.

2.3.1 Response of Output

Figure 2 shows that gross domestic product (GDP), industrial production, and durable goods manufacturing decline significantly more after an oil price increase in the high oil price regime than in the

[14]The impulse responses for the inflation rate in Figure 5 are the standard impulse response functions and not the cumulative impulse response functions.

Figure 2. Responses of Output to an Oil Price Shock



Note: The left-hand column shows the responses to a 9.5 percent oil price increase in the high oil price regime (solid line) and the normal oil price regime (dashed line). The right-hand column shows the difference in the responses (solid line) and their 95 percent confidence bands (dashed lines), where negative values imply that the responses in the high oil price regime are less than the responses in the normal oil price regime.

normal oil price regime. A 9.5 percent oil price increase[15] in the high oil price regime is followed by a cumulative reduction in real GDP of 1.3 percentage points over the next year. In our sample period, real GDP grew on average 2.7 percent per year, so although a 9.5 percent oil price increase probably would not cause a recession, it would be followed by a noticeable slowdown in output growth. In contrast, a 9.5 percent increase in the oil price has a very modest effect on real GDP in the normal oil price regime. Industrial production is a measure of output in manufacturing, mining, and electric and gas utilities. In the high oil price regime, a 9.5 percent oil price increase pushes down industrial production by 2.4 percentage points after one year. A 9.5 percent increase in oil prices, however, only generates a slight increase in industrial production in the normal oil price regime.

A substantial rise in oil prices is expected to affect manufacturing more negatively than the economy as a whole due to manufacturing's greater reliance on energy. High oil prices also could spur a large increase in energy production, which would have a positive effect on industrial production and GDP. For those reasons, we examine the impact of an oil price shock on the manufacturing of durable goods.[16] In the high oil price regime, durable goods manufacturing falls by nearly 3.75 percentage points in the first year after the 9.5 percent oil price increase and continues to decline in year 2. The impulse responses reveal durable goods manufacturing increases by a very small amount after a 9.5 percent rise in oil prices in the normal oil price regime.

2.3.2 Labor Market Variables

Blanchard and Gali (2010) find that oil price shocks have a smaller effect on output when wages are flexible than when wages are sticky

[15] Recall, the average quarterly oil price increase in the high oil price regime is 9.5 percent.

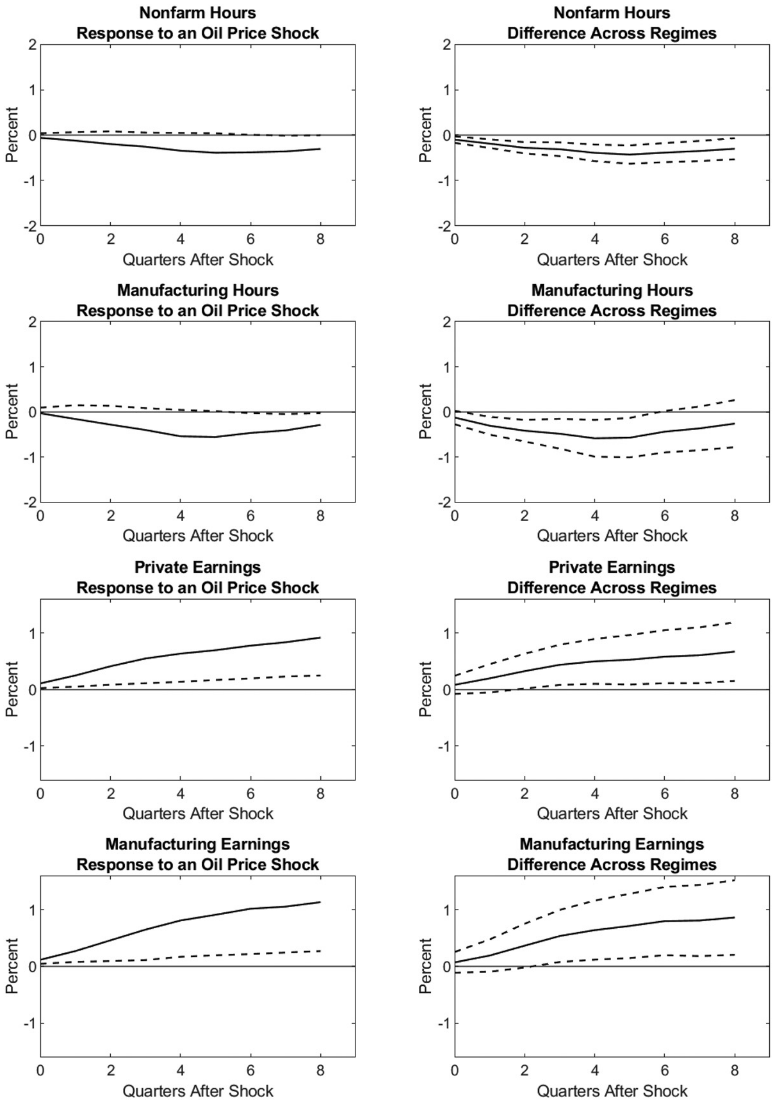
[16] The manufacturing of durable goods represented 38 percent of total industrial production in 2012. According to Federal Reserve Board data release notes, durable goods manufacturing includes the following categories of production: wood product; non-metallic mineral product; primary metal; fabricated metal product; machinery; computer and electronic product; electrical equipment, appliance, and component; motor vehicles and parts; aerospace and miscellaneous transportation equipment; furniture and related product; and miscellaneous.

or rigid. Specifically, an economy with flexible wages produces a larger drop in the real wage rate after an oil price increase, which puts more downward pressure on the real marginal cost. That downward pressure mitigates some of the rise in the marginal cost caused by the higher oil price and, as a result, limits the decline in output. Figure 3 shows the impact of an oil price shock on weekly hours worked (non-farm hours and manufacturing hours) and hourly nominal wages (private earnings and the manufacturing earnings). Non-farm hours and manufacturing hours fall by nearly 0.26 and 0.40 percentage point, respectively, in the first year after a 9.5 percent oil price increase in the high oil price regime, which are significantly greater than their responses in the normal oil price regime. That same 9.5 percent oil price increase also pushes up private earnings and manufacturing earnings in the first year by 0.55 and 0.65 percentage point, respectively, in the high oil price regime, but those same variables increase by 0.11 percentage point in the normal oil price regime, which is consistent with Blanchard and Gali's theoretical model.

A discrepancy exists between our empirical model and the theoretical model presented in Section 3. The nominal wage initially rises and then remains flat for several periods in the theoretical model, but the nominal wage rises continuously in the empirical model. The difference is explained by the fact that our empirical model is estimated over a period with both substantial nominal wage indexation to inflation (the 1970s) and little nominal wage indexation to inflation (post-1982).[17] When wages are indexed to inflation, higher inflation caused by an oil price increase automatically pushes up the lower bound on the nominal wage. That higher lower bound leads to an upward drift in nominal wages, which is consistent with our empirical results. In the absence of wage indexation, inflation does not automatically raise the lower bound on nominal wages. When our empirical model is estimated starting in 1983:Q1 instead of 1972:Q1, the one-year increases in private earnings and manufacturing earnings are only 0.17 and 0.09 percentage point, respectively, in the high oil price regime, while both one-year responses are

[17]Using an estimated New Keynesian model, Keen and Koenig (2018) find that the degree of wage indexation to inflation in the United States is much higher in the 1960s and 1970s than in the post-1982 period.

Figure 3. Responses of Labor Market Variables to an Oil Price Shock



Note: The left-hand column shows the responses to a 9.5 percent oil price increase in the high oil price regime (solid line) and the normal oil price regime (dashed line). The right-hand column shows the difference in the responses (solid line) and their 95 percent confidence bands (dashed lines), where negative values imply that the responses in the high oil price regime are less than the responses in the normal oil price regime.

essentially zero in the normal price regime.[18] Those results from the post-1982 sample are consistent with the findings from our theoretical model, which is calibrated to values consistent with the Great Moderation (1983–2007).

The model presented in Section 3 assigns a key role to labor market rigidities like in Blanchard and Gali (2010). An important distinction of our model, however, is the focus on downward rigid nominal wages as an explanation for the asymmetric responses generated in the high oil price regime. Blanchard and Gali assume that real wages are sticky, which result in symmetric responses to oil price shocks.

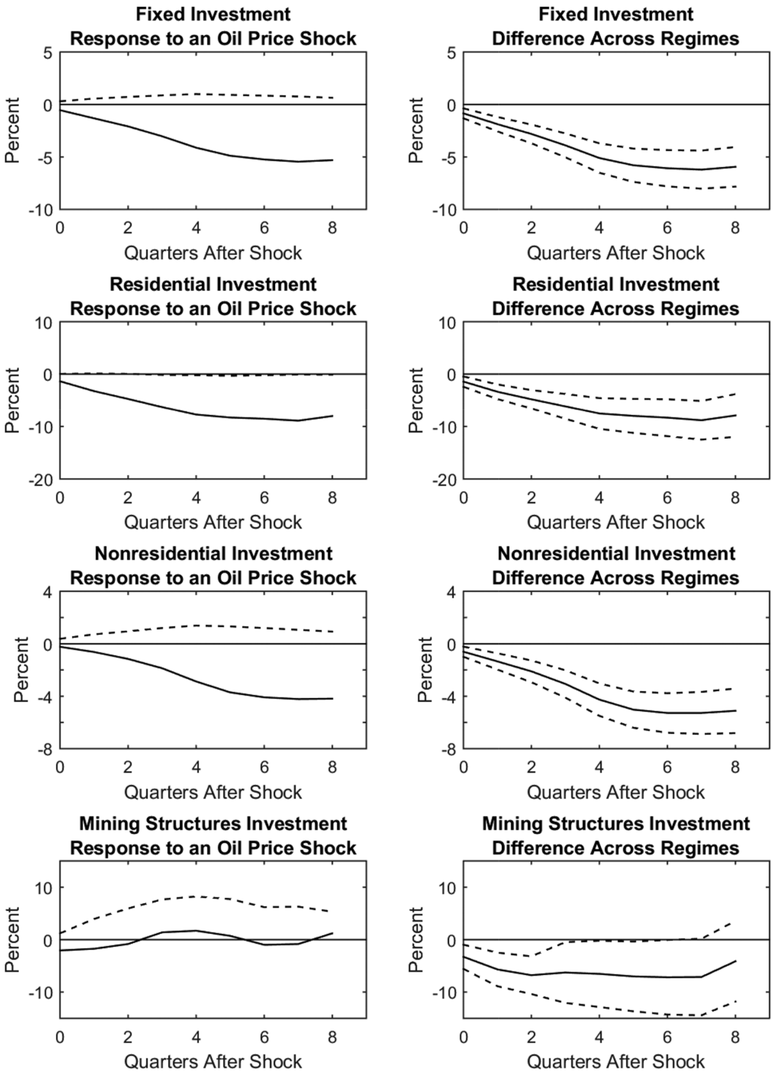
2.3.3 Investment

Figure 4 presents the impact of an oil price shock on private fixed investment, residential investment, non-residential investment, and investment in mining structures. A 9.5 percent oil price increase in the high oil price regime produces a significantly larger decline in private fixed investment, residential investment, and non-residential investment than in the normal oil price regime. Private fixed investment falls more than 4 percentage points in the year after a 9.5 percent oil price increase in the high oil price regime, but it slightly rises in the normal oil price regime. That slight increase in private fixed investment after an oil price increase is caused by the fact that higher oil prices usually stimulate energy-related investment.

Residential investment declines much more than aggregate investment after an oil price increase in the high oil price regime. A 9.5 percent oil price increase in the high oil price regime pushes down residential investment by nearly 8 percentage points over the next year, while the same-sized shock causes residential investment to remain essentially unchanged in the normal oil price regime. Non-residential investment declines by about 2.9 percentage points in the year after a 9.5 percent oil price increase in the high oil price regime, with much of that drop due to a large fall in equipment investment. In the normal oil price regime, the same oil price shock has a small positive effect on non-residential investment. That 9.5 percent oil

[18] The impulse response functions estimated with data starting in 1983:Q1 are included in the appendix, which is available on the authors' websites.

Figure 4. Responses of Investment to an Oil Price Shock



Note: The left-hand column shows the responses to a 9.5 percent oil price increase in the high oil price regime (solid line) and the normal oil price regime (dashed line). The right-hand column shows the difference in the responses (solid line) and their 95 percent confidence bands (dashed lines), where negative values imply that the responses in the high oil price regime are less than the responses in the normal oil price regime.

price increase also pushes up investment in mining structures by 1.7 percentage points in the high oil price regime and by over 8 percentage points in the normal oil price regime.

2.3.4 Consumption and Inflation

Figure 5 displays the impulse response functions for the high oil price and normal oil price regimes and the difference functions for real personal consumption expenditures (PCE), real core PCE, the PCE inflation rate, and the core PCE inflation rate.[19] A 9.5 percent oil price increase in the high oil price regime causes real PCE and real core PCE to decline by over 1 percentage point in the first year. In contrast, a 9.5 percent increase in oil prices has much more moderate effects on real PCE and real core PCE in the normal oil price regime. As for inflation, PCE inflation initially jumps by 0.32 percent and then gradually declines after a 9.5 percent oil price increase in the high oil price regime. The same shock causes the core PCE inflation rate to rise initially by 0.13 percent in the high oil price regime. That number remains steady for about a year and then slowly declines. The 9.5 percent oil price increase has little effect on either inflation measure in the normal oil price regime. The difference in the responses of PCE inflation and core PCE inflation across the two regimes is not statistically significant.

3. Theoretical Model

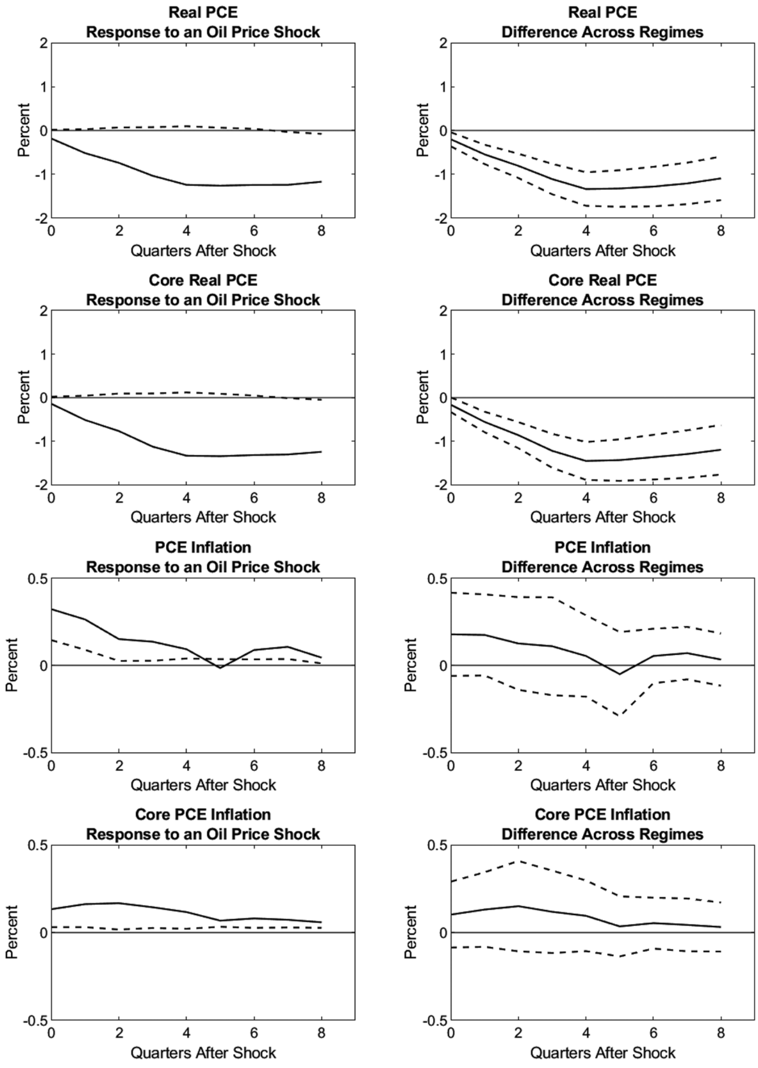
This section develops a New Keynesian model with price stickiness and downward rigid nominal wages to examine the asymmetric effects of key variables to an energy price shock. Price setting follows a Calvo (1983) model of random adjustment, where nominal wages are perfectly flexible on the upside but rigid on the downside. Energy is demanded by households as a consumption good and by firms as a factor of production.[20] The energy endowment each period

---

[19]The impulse response functions and difference functions are cumulative for real PCE and real core PCE, but they are standard (non-cumulative) for PCE inflation and core PCE inflation.

[20]Since energy enters both the households' utility function and the firms' production functions, an energy price shock is equivalent to some combination of

Figure 5. Responses of Consumption and Inflation to an Oil Price Shock



Note: The left-hand column shows the responses to a 9.5 percent oil price increase in the high oil price regime (solid line) and the normal oil price regime (dashed line). The right-hand column shows the difference in the responses (solid line) and their 95 percent confidence bands (dashed lines), where negative values imply that the responses in the high oil price regime are less than the responses in the normal oil price regime.

is sufficient to meet market demand at its exogenously determined price.

The downward nominal wage rigidity is the critical feature in our model that enables energy price shocks to have asymmetric effects. Specifically, a large energy price increase puts downward pressure on real wages, but the downward nominal wage rigidity prevents the nominal wage from falling. That constraint forces firms to reduce both their labor demand and output further, which causes key variables to respond asymmetrically.

Gottschalk (2005), Barattieri, Basu, and Gottschalk (2014), and Hazell and Taska (2018) find evidence in U.S. data that nominal wages are downward rigid. Our model, however, specifically assumes the downward nominal wage rigidity is more likely to bind after a large jump in energy prices. Over the last 40 to 50 years, large oil price increases are considered the primary cause of several U.S. recessions. Daly and Hobijn (2014) and Jo (2021) find evidence that U.S. nominal wages were more rigid during and immediately after those oil-price induced recessions.[21] Their findings support our conjecture that a large energy price increase raises the likelihood the downward nominal wage constraint binds.

3.1 Households

Households are infinitely lived agents who prefer consumption, c_t, but dislike labor, n_t. Each period, households maximize their utility,

a demand shock and a supply shock in a traditional three-equation New Keynesian model. Thus, an energy shock in our model differs from Kim and Ruge-Murcia (2009, 2011), Abbritti and Fahr (2013), and Abo-Zaid (2013), where those papers examine how the downward nominal wage rigidity impacts the asymmetric responses of economic variables to a simple supply shock.

[21] Using Current Population Survey (CPS) data for 1986–2014, Daly and Hobijn (2014) find that the share of workers not receiving a nominal wage increase in a particular year rose during and immediately after the 1991 recession. The Federal Reserve Bank of San Francisco's Wage Rigidity Meter finds similar results with CPS data for the 1981–82 and 1991 recessions. Similarly, Jo (2021) finds that the share of workers not receiving a wage increase rises after the 1980, 1981–82, and 1991 recessions using both the CPS and the Survey of Income and Program Participants data. We should note that Daly and Hobijn, Jo, and the San Francisco Fed's Wage Rigidity Meter also find that the share of workers not receiving a wage increase rises after the 2001 and 2008 recessions, but most economists do not believe oil price increases had a major role in those recessions.

$$U = E_t \sum_{j=0}^{\infty} \beta^j \left[\ln\left(c_{t+j} - \phi_c h_{t+j}\right) - \phi_n \frac{n_{t+j}^{1+\zeta} - 1}{1+\zeta} \right], \tag{2}$$

subject to a consumption aggregator, budget constraint, capital equation, and a nominal wage rigidity that prevents the nominal wage from falling. E_t is the expectational operator at time t, $0 \leq \beta < 1$ is the discount factor, $0 \leq \phi_c < 1$ is the external habit persistence parameter, h_t is the habit persistence variable that is equal to lagged aggregate consumption ($h_t = c_{t-1}$), $\zeta \geq 0$ is the labor supply elasticity, and $\phi_n > 0$. Aggregate consumption is a constant elasticity of substitution (CES) composite of energy consumption, $e_{h,t}$, and non-energy consumption, $c_{n,t}$,

$$c_t = \left(a_1 e_{h,t}^{\upsilon_h} + a_2 c_{n,t}^{\upsilon_h} \right)^{1/\upsilon_h}, \tag{3}$$

where $1/(1-\upsilon_h)$ is the elasticity of substitution between non-energy and energy consumption, and $a_1 > 0$ and $a_2 > 0$ are calibrated such that $a_1(e_{h,t}/c_t)^{\upsilon_h}$ and $a_2(c_{n,t}/c_t)^{\upsilon_h}$ are set equal to energy's and non-energy's shares of consumption, respectively. As a result, the aggregate price level (headline price level), P_t, is a function of the price of non-energy output (core price level), $P_{n,t}$, and the price of energy, $P_{e,t}$:

$$P_t = \left(a_1^{1/(1-\upsilon_h)} P_{e,t}^{\upsilon_h/(\upsilon_h-1)} + a_2^{1/(1-\upsilon_h)} P_{n,t}^{\upsilon_h/(\upsilon_h-1)} \right)^{(\upsilon_h-1)/\upsilon_h}. \tag{4}$$

The households' budget constraint shows the real value of inflows and outflows of funds:

$$\begin{aligned} &\left(\frac{P_{n,t}}{P_t}\right)(c_{n,t} + i_t) + \left(\frac{P_{e,t}}{P_t}\right) e_{h,t} + b_t \\ &\quad = \frac{R_{t-1} b_{t-1}}{\pi_t} + d_t + w_t n_t + \left(\frac{P_{n,t}}{P_t}\right) q_t k_t. \end{aligned} \tag{5}$$

At the beginning of each period, households receive real income from last period's bond holdings, $R_{t-1}b_{t-1}/\pi_t$, where R_t is the gross nominal interest rate between periods t and $t+1$, π_t is the gross headline inflation rate between periods $t-1$ and t, and b_t is the real value of bond holdings. Households then receive their labor income, $w_t n_t$, capital income, $(P_{n,t}/P_t)q_t k_t$, and share of profits from firms and the energy sector, d_t, where w_t is the real wage rate, q_t is the real rental rate of capital, and k_t is capital. Households use those funds to purchase non-energy consumption goods, $(P_{n,t}/P_t)c_{n,t}$, investment goods, $(P_{n,t}/P_t)i_t$, energy, $(P_{e,t}/P_t)e_{h,t}$, and bond holdings, b_t, where i_t is real investment.

Households invest in capital and rent it to the firms in a perfectly competitive market. Once investment decisions are made, capital evolves as follows:

$$k_{t+1} - k_t = i_t \left(1 - S\left(\frac{i_t}{i_{t-1}}\right)\right) - \delta k_t, \tag{6}$$

where $S(\cdot)$ is an investment adjustment cost function that represents the resources lost in the conversion of investment to capital. Following Christiano, Eichenbaum, and Evans (2005), we assume $S(1) = S'(1) = 0$ and $\kappa = S''(1) > 0$.

Households supply labor in a perfectly competitive market but will not accept a nominal wage below its previous level. Thus, we have the following inequality constraint:

$$P_t w_t \geq \gamma P_{t-1} w_{t-1}, \tag{7}$$

where $\gamma \geq 0$ measures the degree of downward nominal wage rigidity. Nominal wages are absolutely downward rigid when $\gamma \geq 1$ but are perfectly flexible when $\gamma = 0$. During the periods when (6) binds, households supply more labor than demanded, so the households' first-order condition for labor does not bind.

3.2 Firms

Firms are monopolistically competitive producers of non-energy output, $y_{n,t}$. Firm f uses its inputs of capital, $k_{f,t}$, labor, $n_{f,t}$, and energy, $e_{f,t}$, to produce its output, $y_{f,t}$, according to the following production function:

$$y_{f,t} = \left(b k_{f,t}^{\upsilon_f} + (1-b) e_{f,t}^{\upsilon_f}\right)^{\alpha/\upsilon_f} (n_{f,t})^{1-\alpha}, \tag{8}$$

where $1/(1-\upsilon_f)$ is the elasticity of substitution between energy and capital, $0 < b < 1$, and $0 < \alpha < 1$. The capital and labor used by firm f are rented for the nominal capital rental rate of $P_{n,t}q_t$ and the nominal wage rate of $P_t w_t$, respectively. Firm f also purchases its energy input in a perfectly competitive market for a price of $P_{e,t}$. Given those capital, labor, and energy costs, firm f minimizes its production costs, $P_{n,t}q_t k_{f,t} + P_t w_t n_{f,t} + P_{e,t}e_{f,t}$, subject to (7) to generate its input factor demands.

The differentiated output, $y_{f,t}$, produced by a continuum of firms ($f \in [0,1]$) are combined to generate aggregate non-energy output, y_t, using the Dixit and Stiglitz (1977) method:

$$y_t = \left[\int_0^1 y_{f,t}^{(\epsilon-1)/\epsilon} df\right]^{\epsilon/(\epsilon-1)}, \tag{9}$$

where $-\epsilon$ is the price elasticity of demand for $y_{f,t}$. Since firm f sells $y_{f,t}$ at a price of $P_{f,t}$, cost minimization by households means the demand for $y_{f,t}$ is a decreasing function of its relative price:

$$y_{f,t} = \left(\frac{P_{f,t}}{P_{n,t}}\right)^{-\epsilon} y_t, \tag{10}$$

where $P_{n,t}$ is a non-linear price index of a continuum of non-energy output:

$$P_{n,t} = \left[\int_0^1 P_{f,t}^{(1-\epsilon)} df\right]^{1/(1-\epsilon)}. \tag{11}$$

Non-energy output comprises non-energy consumption and investment, $y_t = c_{n,t} + i_t$.

Price setting follows the Calvo (1983) model of random adjustment. Each period, a fraction of firms, $(1-\eta)$, can optimally readjust their prices, while the remaining fraction, η, raise their prices by last period's core inflation rate, $\pi_{n,t-1}$. When presented with a price adjustment opportunity, firm f selects a price, $P^*_{f,t}$, that maximizes the present value of current and expected future profits given the probability of future adjustment opportunities:

$$\max_{P^*_{f,t}} E_t \left[\sum_{j=0}^{\infty} \beta^j \eta^j \lambda_{t+j} \left(\frac{\Pi_{n,t+j} P^*_{f,t}}{P_{t+j}} y_{f,t+j} - w_{t+j} n_{f,t+j} \right.\right.$$
$$\left.\left. - \frac{P_{n,t+j}}{P_{t+j}} q_{t+j} k_{f,t+j} - \frac{P_{e,t+j}}{P_{t+j}} e_{y,t+j} \right) \right], \tag{12}$$

where

$$\Pi_{n,t+j} = \left\{ \begin{array}{ll} \pi_{n,t} \times \pi_{n,t+1} \times \cdots \times \pi_{n,t+j-1} & \text{for } j \geq 1 \\ 1 & \text{for } j = 0 \end{array} \right\} \tag{13}$$

subject to the demand for its product, (9), and its input factor demands.

3.3 Energy

Energy is used by households as a consumption good and by firms as a factor input. Therefore, aggregate energy, e_t, comprises energy consumed by both households and firms:

$$e_t = e_{h,t} + e_{f,t}. \tag{14}$$

The energy endowment is sufficient to meet market demand at an exogenously determined price. As in Wei (2003), the real price of energy, $p_{e,t} = P_{e,t}/P_t$, follows an AR(1) process:

$$\ln(p_{e,t}) = \rho_e \ln(p_{e,t-1}) + \varepsilon_t, \tag{15}$$

where $0 \leq \rho_e < 1$ and $\varepsilon_t \sim N(0, \sigma_e)$.[22]

3.4 Monetary Policy

Monetary policy is conducted via a Taylor (1993) style nominal interest rate rule with interest rate smoothing. That is, the central bank adjusts its nominal interest rate target, R_t, in response to changes in

[22] Our qualitative results are the same if the energy price is assumed to follow an ARMA(1,1) process, or if the model is solved with the quantity of energy, rather than the price of energy, being exogenous.

the lagged nominal interest rate, R_{t-1}, the core inflation rate, $\pi_{n,t}$, and non-energy output, y_t:

$$\begin{aligned}\ln(R_t/R) &= \theta_R \ln(R_{t-1}/R) \\ &\quad + (1-\theta_R)[\theta_\pi \ln(\pi_{n,t}/\pi_n^*) + \theta_y \ln(y_t/y_t^P)],\end{aligned} \tag{16}$$

where π_n^* is the gross steady-state core inflation rate, y_t^P is potential non-energy output, $0 \leq \theta_R < 1$, $\theta_\pi > 1$, and $\theta_y \geq 0$. Potential non-energy output is the level of non-energy output that would exist in the absence of nominal price and wage frictions.

4. Equilibrium and Calibration

Our model's systematic equilibrium encompasses the set of difference equations representing the model's first-order conditions, the identity equations, and the exogenous energy price shock process. The long-run trend in the core price level, the headline price level, and the price of energy means that all of the nominal variables, except R_t, must be divided by P_t to induce stationarity.[23] Our system of equations is linearized around its steady state, and the standard solution techniques (e.g., see Sims 2002) are utilized to find the rational expectations solution. Finally, the Holden and Paetz (2012) algorithm is used to simulate our linear New Keynesian model with a downward rigid nominal wage inequality constraint.[24]

Table 2 displays the parameters calibrated to quarterly values. To begin, the discount factor, β, is parameterized to 0.99; the degree of habit persistence, ϕ_c, is set to 0.7; the degree of downward nominal wage rigidity, γ, equals 1; and the preference parameter, ϕ_n, is calibrated so the steady-state level of labor, n^*, equals 0.3. The

[23]We assume the core price level, the headline price level, and the price of energy all have identical long-run trends, so energy's share of the economy remains constant in the long run.

[24]Holden and Paetz (2012) develop a method to solve and simulate New Keynesian models with occasionally binding constraints. In addition to solving the model when the constraint binds, the authors' algorithm uses a hybrid local/global approximation to account for the possibility the constraint will bind in the future, even when the constraint is not currently binding.

Table 2. Calibrated Parameter Values

Parameter	Symbol	Value
Discount Factor	β	0.99
Habit Persistence in Consumption	ϕ_c	0.7
Degree of Downward Nominal Wage Rigidity	γ	1
Steady-State Labor	n^*	0.3
Frisch Labor Supply Elasticity	$1/\zeta$	0.72
Price Elasticity of Demand	ϵ	6
Depreciation Rate	δ	0.025
Investment Adjustment Costs Parameter	κ	2.5
Capital and Energy's Share in Production	α	0.33
Probability of Non-optimal Price Adjustment	η	0.75
CES Consumption/Energy Substitution Parameter	ν_h	–0.9
CES Capital/Energy Substitution Parameter	ν_f	–0.9
Energy's Share Used in Consumption	e_h/c	0.060
Energy's Share Used in Production	b	0.045
Monetary Policy's Reaction to Inflation	θ_π	1.5
Monetary Policy's Reaction to Output	θ_y	0.125
Monetary Policy's Reaction to Lagged Nominal Rate	θ_R	0.7
Steady-State Gross Core Inflation Rate	π_n^*	1.005
AR Coefficient in the Energy Price Shock	ρ_e	0.95

Frisch labor supply elasticity, $1/\zeta$, is fixed to Heathcote, Storesletten, and Violante's (2014) estimate of 0.72.[25] The values of a_1 and a_2 from the aggregate consumption equation, (3), are set so the ratio of energy used in consumption to aggregate consumption equals its average of 0.060 from 1972:Q1 to 2017:Q4.[26] The parameter υ_h used to calculate the elasticity of substitution between energy and non-energy consumption, $1/(1-\upsilon_h)$, equals –0.9. That value used by Gavin, Keen, and Kydland (2015) implies that the two goods are compliments. We assume that the price elasticity of demand, ϵ, is 6, so the steady-state markup of price over marginal cost is 20 percent. The capital depreciation rate, δ, is calibrated to 0.025, while

[25]Heathcote, Storesletten, and Violante (2014) estimate the Frisch labor supply elasticity to be 0.72 when a household is defined as a husband and a wife. Given that many New Keynesian models utilize higher values for the Frisch labor supply elasticity, we examine the sensitivity of our results to higher values later in the paper.

[26]The ratio of energy consumption to aggregate consumption is calculated as the average ratio of personal consumption expenditures: goods and services to personal consumption expenditures.

the investment adjustment costs parameter, κ, is set to Christiano, Eichenbaum, and Evans' (2005) estimate of 2.5.

In the production function, capital and energy's share in production, α, is set to 0.33, while b is fixed so energy's share in production equals its average of 0.045 from 1972:Q1 to 2017:Q4.[27] We follow Gavin, Keen, and Kydland (2015) and parameterize υ_f to –0.9. Since $\upsilon_f < 0$, capital and energy are complimentary goods. The Calvo (1983) probability of non-optimal price adjustment, η, is calibrated to 0.75, which implies that a firm on average optimally readjusts its price once a year. The steady-state relative prices of energy and non-energy, P_e and P_n, are assumed to be equal. As for the policy rule, the parameters on inflation and output, θ_π and θ_y, are calibrated to Taylor's (1993) estimates of 1.5 and 0.125, respectively, while the coefficient on the lagged interest rate, θ_R, is fixed to 0.7.[28] The gross steady-state quarterly core inflation rate, π_n^*, is equal to 1.005, which is consistent with a 2 percent annual inflation rate target.[29] Finally, the AR coefficient in the energy price shock process, ρ_e, is set to 0.95 as in Wei (2003).

5. Model Results

5.1 Impulse Response Functions

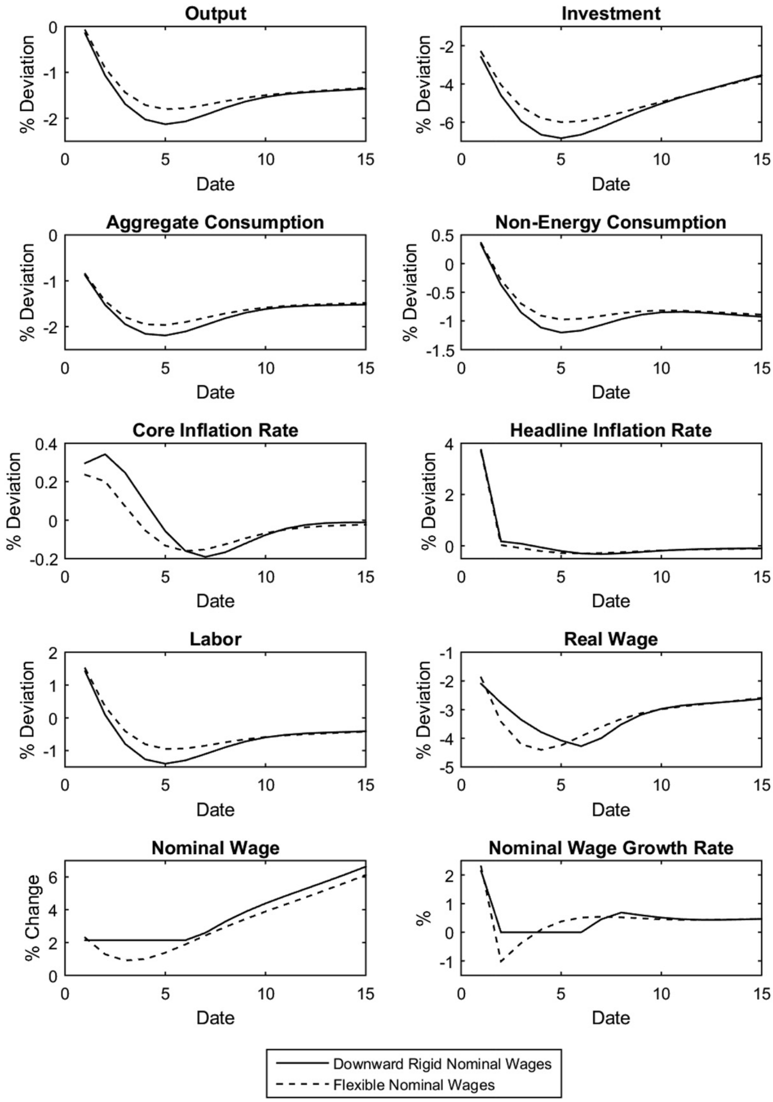
Figure 6 shows the impulse response functions of key variables to a 35 percent increase in the energy price both with flexible nominal wages ($\gamma = 0$) and downward rigid nominal wages ($\gamma = 1$). A 35 percent energy price increase approximately matches the large rise in U.S. energy prices during 1974:Q4, 1990:Q3, and 1990:Q4. The impulse responses for output, investment, aggregate consumption, non-energy consumption, core inflation rate, headline inflation,

[27]The 2020 Annual Energy Review publishes annual data on energy's share of GDP (see Table 1.7). The annual level of energy's share of production is calculated by subtracting personal consumption expenditures: goods and services share of GDP from energy's share of GDP. Finally, our parameter is calibrated to the average level of energy's share of GDP from 1972 to 2017.

[28]Since the model is quarterly, the Taylor (1993) rule coefficient on output, θ_y, of 0.5 is divided by 4.

[29]The steady-state core inflation rate, π_n^*, equals the steady-state headline inflation rate, π^*, in our model.

Figure 6. Impulse Responses to a 35 Percent Oil Price Increase



labor, and the real wage are the percent deviations from their steady states, while the responses for the nominal wage and nominal wage growth rate are the percentage change from its initial value and the actual growth rate, respectively.[30] In a linearized model, identically sized increases and decreases in energy prices have symmetric effects. The presence of downward rigid nominal wages, however, causes the effects of an energy price increase to differ in magnitude from the effects of the same-sized energy price decline. The difference between the impulse responses in Figure 6 illustrates the approximate asymmetric effects of an energy price shock in our model with downward rigid nominal wages.

The dashed line in Figure 6 reveals the impact of a rise in energy prices in a model with flexible nominal wages. In that model, an energy price increase immediately pushes up production costs, which causes firms to reduce their supplies of output and to raise prices. The lower supply of output puts downward pressure on real wages and capital rents by reducing firms' demand for labor and capital. Households respond to higher energy prices and smaller capital rents by reducing their demand for energy consumption and investment. Higher energy prices also reduce the relative price of non-energy goods, so households substitute some of their lost energy consumption for additional non-energy consumption to accommodate their preferences for habit persistence in aggregate consumption. That shift moderates the declines in aggregate consumption and non-energy output after an energy price increase. In the labor market, firms demand less labor, but households respond to the decline in aggregate consumption by increasing their labor supply and decreasing their leisure.[31] The increase in labor supply combined with the decrease in labor demand pushes down the real wage, but it pushes up labor hours. The initial jump in headline inflation, however, is large enough to dominate the fall in the real wage, so the nominal wage rises.

[30] In the steady state, prices rise by 0.5 percent each period. Thus, the steady-state nominal wage increases by 0.5 percent each period, which means the steady-state nominal wage growth rate is a constant 0.5 percent.

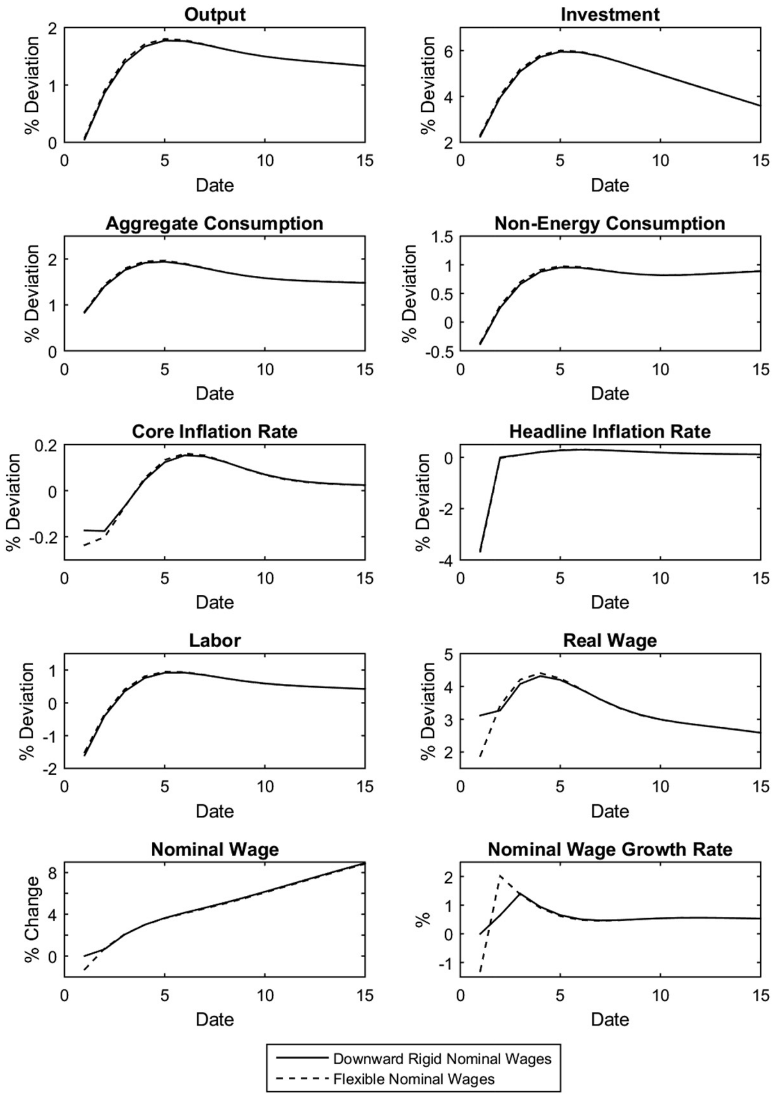
[31] Bodenstein, Guerrieri, and Gust (2013) argue that households do not increase labor supply after a negative energy shock when a Greenwood, Hercowitz, and Huffman (1988) style of utility function is used instead of the additively separable utility function employed here. One impact from labor supply not rising is that non-energy consumption decreases, rather than increases, after a positive energy price shock.

In subsequent periods, elevated energy prices retreat slowly, which leads to a moderation of inflation. As more firms have an opportunity to raise their prices in response to the energy price increase, output proceeds to fall for several more periods. Furthermore, the slow adjustment of consumption and investment due to habit persistence in consumption and investment adjustment costs means households' demand for non-energy consumption continues to fall in the short term. The continued decline in output demand and supply causes output to fall for another four periods. Reduced production lowers labor demand, which puts downward pressure on the real wage and labor hours. The decline in the real wage dominates the moderation of headline inflation, so the nominal wage growth rate turns negative. After several periods, firms start to lower their prices in response to declining energy prices, which stimulates output and pushes the economy back to its steady state.

The solid line in Figure 6 shows the effects of downward rigid nominal wages on the responses of key economic variables to an energy price increase. The main effects of the downward wage constraint begin to occur in the first period after the energy price increases when the downward nominal wage rigidity prevents the nominal wage from declining. Firms react to those higher labor costs by reducing their output further and by raising their prices more. The price increases cause households to enhance their cuts to aggregate consumption and investment. The effects of the downward nominal wage rigidity continue to affect the economy directly, as long as the nominal wage is higher than it otherwise would have been in the absence of the downward rigidity. Even after the downward wage constraint no longer binds, previous pricing decisions and lower capital investment continue to dampen output for a few more periods relative to the flexible nominal wage specification. The nominal wage growth rate in our model remains at zero for several periods, which indicates the downward nominal wage rigidity is binding. Thus, the downward rigid nominal wage model produces a larger and more persistent output decline than in the flexible nominal wage model.

Figure 7 displays the impulse response functions of key economic variables to a 35 percent decrease in the energy price. The solid line represents responses of the model with downward rigid nominal wages, and the dashed line denotes responses of the model with flexible nominal wages. The key finding from these impulse responses is

Figure 7. Impulse Responses to a 35 Percent Oil Price Decrease



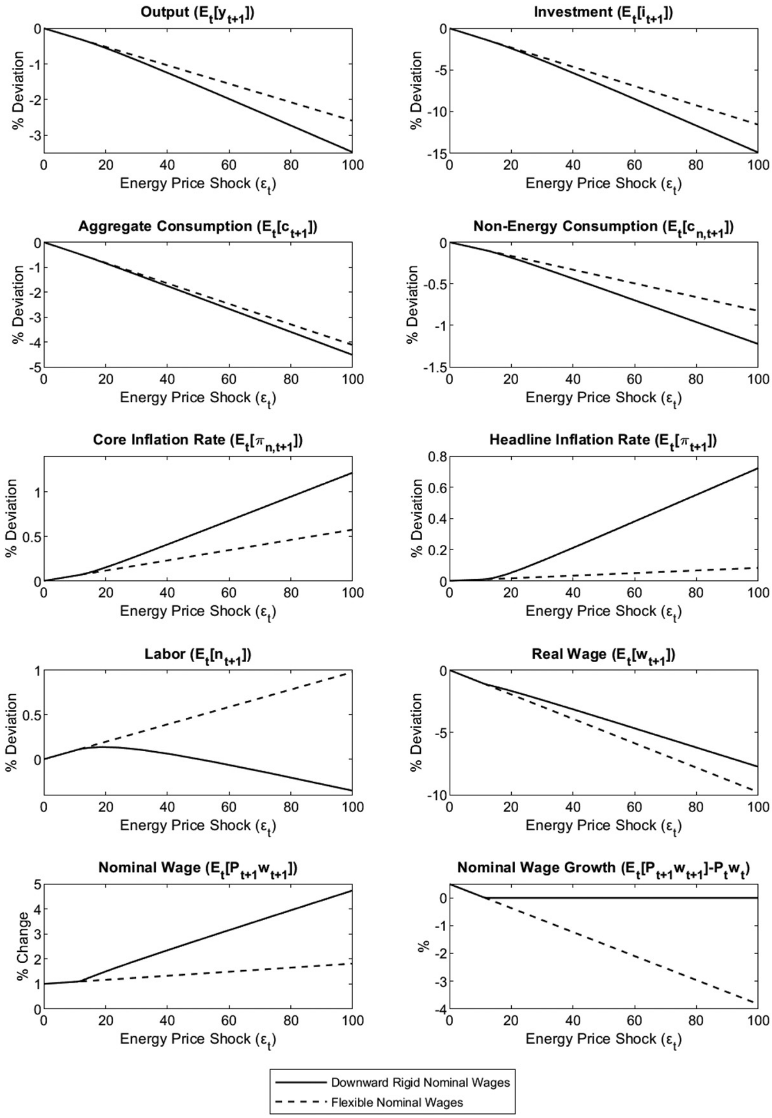
a large fall in energy prices only causes the downward rigid nominal wage constraint to bind in the initial period. The nominal wage constraint binds because lower energy prices push down the headline price level at a faster rate than firms' increased demand for labor drives up the market-clearing real wage. Hence, the actual real wage is above its market-clearing level, which causes price-adjusting firms to limit the decline in their prices leading to a slightly smaller increase in output. Even though the downward rigid wage constraint does not bind in future periods, the output response is slightly lower for a few more periods than in the flexible wage model because price stickiness delays the opportunity for initial price-adjusting firms to adjust their prices again.

Our findings indicate that a large energy price decrease generates impulse response functions for the downward rigid nominal wage model that are very similar to the responses for the flexible nominal wage model, especially in periods $t+1$ and beyond. Since energy price increases and decreases have symmetric effects on the flexible wage model, the downward rigid wage model's asymmetric impulse responses are due primarily to the wage constraint binding after an energy price increase, as opposed to an energy price decrease. Therefore, the remainder of our analysis focuses on comparing the impact of energy price increases on our downward rigid nominal wage model and flexible nominal wage model.

5.2 Decision Rules

Figure 8 presents the period $t+1$ decision rules for key economic variables as a function of an energy price shock, ε_t, that ranges from a 0 percent to 100 percent increase in the price of energy. We focus on the period $t+1$ decision rules because if ε_t is large enough, the downward rigid nominal wage constraint begins to bind in the first period following an energy price increase (i.e., period $t+1$). The solid line displays the impact of an energy price increase on a model with downward rigid nominal wages, while the dashed line shows its effect on a model with flexible nominal wages. Decision rules for the model with flexible wages are linear because the model is solved using standard linearization techniques. The downward rigid nominal wage constraint, however, introduces a non-linear feature into the otherwise linear flexible wage model. Thus, any deviation of the downward rigid nominal wage model's decision rules from the flexible nominal wage model's decision rules represents the asymmetric

Figure 8. Period $t + 1$ Decision Rules as a Function of the Energy Price Shock (ε_t)



and non-linear effects that are attributable to the downward rigid nominal wage constraint.

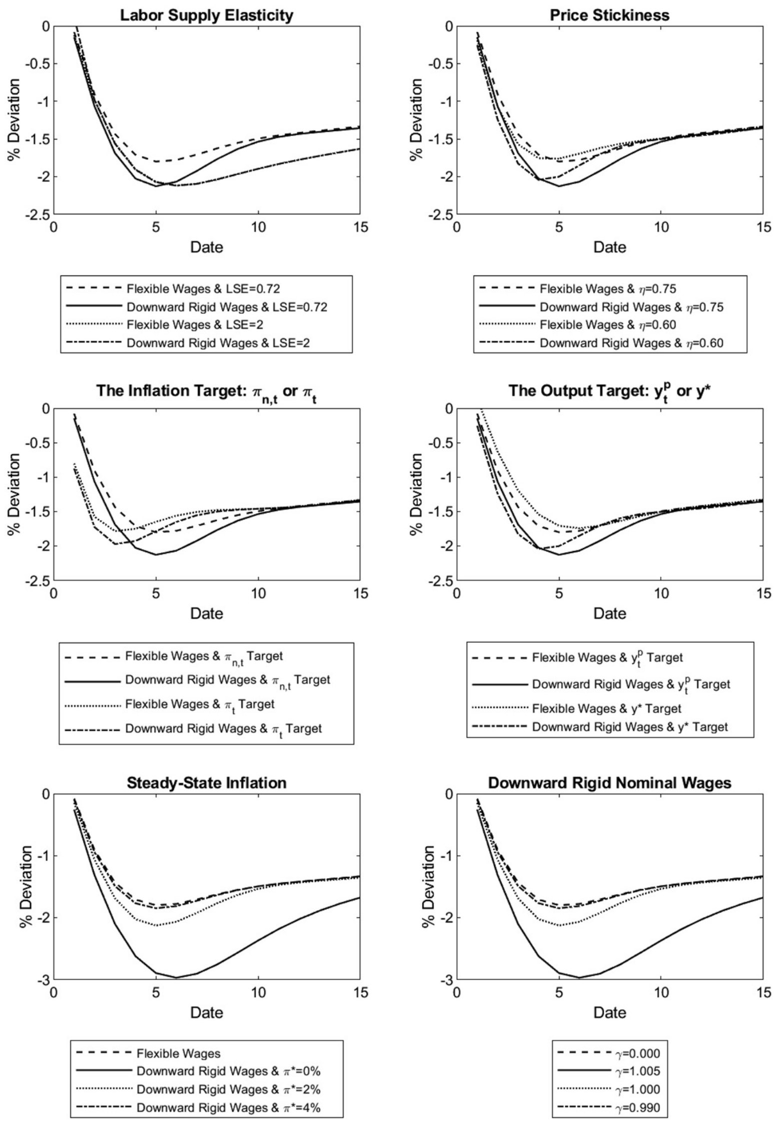
The results in Figure 8 reveal that for small energy price shocks, $\varepsilon_t \leq 11\%$, the nominal wage growth rate is positive, so both models generate identical results because the downward rigid nominal wage constraint does not bind. When $\varepsilon_t > 11\%$, the nominal wage cannot fall in the downward rigid nominal wage model, so output, aggregate consumption, non-energy consumption, and investment decline at faster rates than in the flexible wage model, while core inflation and headline inflation rise at quicker rates. The spread between the solid and dashed lines continues to grow as the size of the oil price shock rises, which indicates the responses from the downward rigid nominal wage model are rising in a non-linear manner. That result suggests a model with downward rigid nominal wages generates responses to an energy price increase consistent with both Hamilton's (1996, 2003, 2011) net oil price increase model and Kilian and Vigfusson's (2013) net oil price change model.

5.3 Sensitivity Analysis

The asymmetric effects of an energy price shock in a New Keynesian model with downward rigid nominal wages depend, sometimes critically, on the calibration of certain parameters. Specifically, output's response to an energy price shock depends on the value of the labor supply elasticity, the degree of price stickiness, the monetary authority's choice of inflation and output targets, the amount of steady-state inflation, and the degree of downward rigid nominal wages. Figure 9 illustrates the impact of those features on output's response to a 35 percent increase in the energy price.

The top, left-hand graph of Figure 9 displays the impact of the labor supply elasticity on output's response to a 35 percent energy price increase. Most models with downward rigid nominal wages assume the labor supply elasticity is very low. For example, Benigno and Ricci (2011) set the labor supply elasticity to 0.4, while Schmitt-Grohe and Uribe (2016) set the labor supply elasticity to 0. Our baseline model calibrates the labor supply elasticity to Heathcote, Storesletten, and Violante's (2014) estimate of 0.72. Since others, such as Christiano and Eichenbaum (1992), use a much higher labor supply elasticity, we also examine output's response

Figure 9. Output's Response to a 35 Percent Oil Price Increase: A Sensitivity Analysis



when the labor elasticity is equal to 2. The solid and dashed lines display the impulse responses for a downward rigid nominal wage model and a flexible nominal wage model, respectively, when the labor supply elasticity (LSE) equals 0.72. A comparison of those impulse responses reveals output falls substantially more when nominal wages are downward rigid. When the labor supply elasticity is set to 2, the dash-dotted line and the dotted line show output's responses are essentually identical in the models with downward rigid nominal wages and flexible nominal wages. That is, output's asymmetric response to an energy price increase in a model with downward rigid nominal wages essentially disappears when labor supply elasticity is 2. The intuition is that a higher labor supply elasticity indicates a flatter labor supply curve. When an energy price increase causes labor demand to decrease, a flatter labor supply curve limits the decline in the real wage. A large drop in the real wage is necessary to offset the inflationary effects of the energy price increase, so the downward rigid nominal wage constraint binds.

The effect of the degree of price stickiness on output's response to an energy price increase is displayed in the top, right-hand graph of Figure 9. The solid and dashed lines show the impact of a 35 percent energy price increase on output in the downward rigid nominal wage and flexible nominal wage models, respectively, when prices change on average once a year, $\eta = 0.75$. We next examine the effect of an energy price increase when prices adjust on average once every 2.5 quarters, $\eta = 0.60$. The dash-dotted line and dotted line illustrate the responses of output in the downward rigid wage model and flexible nominal wage model, respectively, when $\eta = 0.60$. Our results show that a modest reduction in the degree of price stickiness leads to slightly larger responses in the short run, but those responses are not as persistent. In terms of the degree of asymmetry, a higher degree of price stickiness causes the asymmetry in the model to persist for a longer period of time.

The impact of the monetary authority's choice of inflation and output targets on output's response to a 35 percent energy price increase is examined in the middle graphs of Figure 9. The solid and dashed lines on the middle, left-hand graph display output's responses in the downward rigid nominal wage and flexible nominal wage models, respectively, when core inflation, $\pi_{n,t}$, is the target of monetary policy, while the dash-dotted and dotted lines show output's responses in the same models when headline inflation, π_t, is the policy target. The impulse responses for output are less asymmetric in the downward rigid nominal wage model in economies where the monetary authority targets headline inflation. When the energy price jumps, the initial increase in headline inflation is much larger than in core inflation, but in subsequent periods, headline inflation lags behind core inflation as the energy price slowly declines. That higher initial increase causes the nominal interest rate to rise more on impact when monetary policy targets headline inflation, as opposed to core inflation. The elevated nominal rate pushes down output more aggressively, which leads to a larger initial decline in the real wage. In subsequent periods, the real wage, having already experienced a large drop, does not decline as much as it does when core inflation is the policy target. That smaller decrease prevents the downward nominal wage rigidity from binding as long, and as a result, the degree of asymmetry in output's response is more modest when monetary policy targets headline inflation.

A similar but less dramatic change in output's response to a 35 percent energy price increase occurs when the monetary authority targets steady-state output, y^*, as opposed to potential output y_t^p. The solid and dashed lines on the middle, right-hand graph of Figure 9 display output's responses in the downward rigid nominal wage and flexible nominal wage models, respectively, when monetary policy targets potential output, while the dash-dotted and dotted lines show output's responses in the same models when steady-state output is the policy target. The key difference between potential output and steady-state output is that potential output falls after an energy price increase, while steady-state output remains unchanged. As a result, the output gap (actual output minus its target) decreases more when the monetary authority targets steady-state output. That larger decline in the output gap dampens the policy-induced increase in the nominal interest rate, which leads to higher headline and core inflation and a smaller drop in the real wage. The inflation and real wage responses put upward pressure on the nominal wage, which reduces the asymmetry in output's response when steady-state output rather than potential output is the policy target.

The bottom, left-hand graph of Figure 9 illustrates how the steady-state inflation rate, π^*, affects output's response to an energy price shock. The dashed line displays output's response for a flexible nominal wage model. Regardless of the level of the steady-state inflation rate, the impulse response functions for output always remain the same in the flexible wage model. The solid, dotted, and dash-dotted lines represent output's response to an energy price increase in the downward rigid nominal wage model when the steady-state annual inflation rate is 0 percent, 2 percent, and 4 percent, respectively. The differences between each line and the dashed line indicate the degree of asymmetry in output's response to an energy price increase. Those results demonstrate that the degree of asymmetry is the greatest when the steady-state inflation rate is low, 0 percent, and is much more muted when the steady-state inflation rate is high, 4 percent. It follows that when the steady-state inflation rate is higher, the real wage has to fall more before it hits the downward nominal wage rigidity that causes the asymmetric responses after an energy price increase.

The degree of the downward nominal wage rigidity also influences output's response to an energy price increase. The bottom, right-hand graph of Figure 9 shows that output's response becomes more asymmetric as the degree of downward nominal wage rigidity rises. The dashed line displays output's response to an energy price shock when nominal wages do not have a downward constraint ($\gamma = 0$). The solid, dotted, and dash-dotted lines represent output's response when nominal wages must rise by at least 0.5 percent a period ($\gamma = 1.005$), cannot fall at all ($\gamma = 1.000$), and can only fall by 0.5 percent a period ($\gamma = 0.995$), respectively. The differences between each line and the dashed line represent the impact of downward nominal wage rigidity on the asymmetry in output's response to an energy price increase. When nominal wages exhibit a high degree of downward rigidity (i.e., γ is large), an energy price increase is more likely to push down the real wage enough to cause the nominal wage to hit its downward constraint. The sooner the nominal wage bumps into that constraint, the greater the asymmetry in the impulse response functions after an energy price increase.

Our sensitivity analysis reveals that the ability of an energy price increase to generate asymmetric impulse response functions in a model with downward rigid nominal wages depends critically on a few key parameters. Specifically, an energy price shock is more likely to produce asymmetric responses when the labor supply elasticity is low, the degree of price stickiness is large, the monetary authority targets core inflation and potential output, the steady-state inflation rate is low, and the degree of downward rigid nominal wages is high.

5.4 Comparing the 1970s with the 2000s

Oil price shocks produced much larger effects on output in the 1970s than in the 2000s. Compared with the 2000s, the 1970s was a period in which energy's shares of consumption and output were larger, monetary policy responded less aggressively to inflation, the steady-state inflation rate was higher, and wage indexation to inflation was more prevalent. This section examines how those factors affect output's response to an energy price shock in New Keynesian models with and without downward rigid nominal wages.

Our calibration of the 1970s' economy and the 2000s' economy will reflect the difference in energy's shares, the monetary policy rule, steady-state inflation, and the degree of downward rigid nominal wages that existed during those periods. The model of the 1970s is calibrated to data from 1973:Q1 through 1979:Q3, which coincides roughly with the period from the end of the Bretton Woods fixed exchange rate system to the beginning of the Volcker disinflation in October 1979. The model representing the 2000s is calibrated to data from 2000:Q1 through 2017:Q4. Beginning with the monetary policy rule, the parameters, θ_π, θ_y, and θ_R, from the 1970s' model are set to Mehra's (2002) estimates of 1.1, 0.1625, and 0.44, respectively, and their counterparts for the 2000s' model are set to our baseline values of 1.5, 0.125, and 0.7, respectively.[32] Energy's shares of consumption and production are set to their averages of 0.073 and 0.051, respectively, in the 1970s' model and to 0.056 and 0.040, respectively, in the 2000s' model. The average annual inflation rate was 8 percent in the 1970s, so π_n^* is set to 1.02 in the 1970s' model. During the 2000s, the Federal Reserve targeted a 2 percent inflation rate, so π_n^* is set to 1.005 in that model.[33] Holland (1988, 1995) shows that the higher inflation rates of the 1970s led to many workers, both unionized and non-unionized, receiving automatic cost-of-living increases. Those automatic raises caused the degree of the downward nominal wage rigidity, γ, to increase. Blanchard and Gali (2010) acknowledge

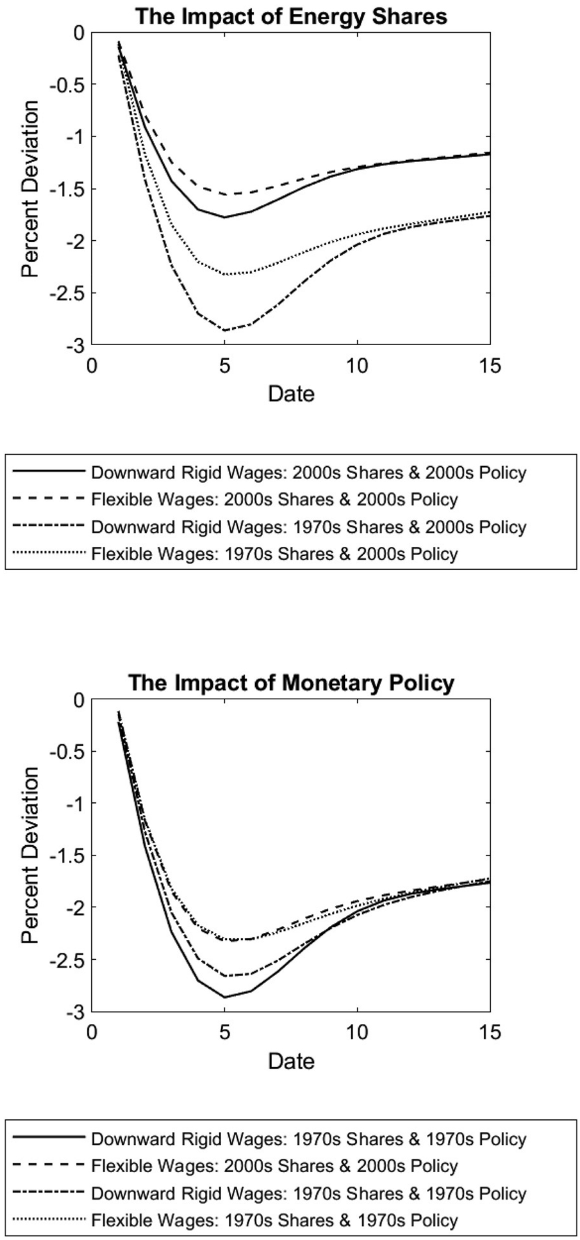
[32]Like the baseline model, the θ_y coefficients of 0.65 and 0.5 are divided by 4 because the model is quarterly.

[33]In the early 2000s, the market participants believed the Federal Reserve was targeting an inflation rate of 2 percent. That 2 percent inflation target was confirmed by the Federal Reserve on January 25, 2012.

this connection by stating, "the 1970s were times of strong unions and high wage indexing" to inflation. Since we do not have a precise estimate for γ, we calibrate the parameter to 1.015 in the 1970s' model, but keep it equal to the baseline value of 1.000 in the 2000s' model. Those values reflect the higher level of wage indexation in the 1970s, while maintaining the spread between π_n^* and γ in the 1970s' model and 2000s' model, so our analysis can focus on the impact of differences in energy's shares and monetary policy on the two models.

Figure 10 compares the impact of an energy price shock on output in the 1970s, when energy's shares of consumption and production were larger and monetary policy was less focused on inflation, with that of the 2000s. The top graph of Figure 10 displays the effect of the different energy's shares on output's response to a 35 percent energy price increase when the monetary authority follows a 2000s' policy. The solid and dashed lines represent output's response with energy shares from the 2000s in both the downward rigid nominal wage model and the flexible wage model, respectively, while dash-dotted and dotted lines show the same respective responses with energy shares from the 1970s. Those impulse response functions reveal that the larger energy shares in the 1970s enhance the decline in output after a 35 percent energy price increase in both models. The bottom graph of Figure 10 shows how the change in monetary policy from the 1970s to the 2000s affects output's response after an energy price shock when energy's shares are at 1970s' levels. The solid and dashed lines represent output's response with monetary policy from the 2000s in both the downward rigid nominal wage model and the flexible wage model, respectively, while dash-dotted and dotted lines show the same respective responses with monetary policy from the 1970s. The impulse responses show that the change in monetary policy has no meaningful effect on output's response in the flexible wage model. Output's response, however, is more muted in the downward rigid wage model with 1970s' monetary policy than with 2000s' monetary policy. The larger emphasis on output stability and the smaller emphasis on inflation stability in the 1970s' monetary policy caused the real wage to fall less and inflation to rise more, which put upward pressure on the nominal wage after an energy price increase. Therefore, our downward rigid nominal wage model attributes the larger output decline after an energy price increase in the

Figure 10. Output's Response to a 35 Percent Oil Price Increase: A Comparison of the 1970s with the 2000s



1970s to that period's higher level of energy's shares of consumption and production and not to that period's monetary policy.[34]

5.5 Demand or Supply Shock: Does It Matter?

Our model assumes an energy price shock is exogenous to the economy, and the energy endowment is sufficient to meet energy demand at that price. Since the energy price does not respond to changes in the domestic economy, our model views the energy price shock as a disturbance that originates internationally. A sampling of the largest energy price shocks over our estimation period is consistent with that assumption. Recall from the discussion of Figure 1 that the large oil price increases of 1973–74, 1979–80, 1981, and 1990 are usually attributed to foreign oil supply disruptions, while the 2002–08 oil price spike is often attributed to the large rise in oil demand from China and India.

One drawback of our model is that it does not distinguish between energy price shocks caused by changes in foreign energy demand and supply. We can use an open-economy New Keynesian model by Balke and Brown (2018), however, to infer how energy price shocks caused by changes in foreign energy demand and supply affect our model. Specifically, the authors examine separately the impact of a rise in energy prices caused by an increase in foreign demand or a decrease in foreign supply. They find that an increase in foreign energy demand generates a larger rise in domestic prices and a smaller decline in domestic output than a comparable fall in foreign energy supply. Balke and Brown's rationale is straightforward in that an increase in foreign energy demand is usually caused by a growing foreign economy that is demanding and producing more goods and services. Some of the increased foreign demand for goods and services is produced in the domestic country. Therefore, a rise in energy prices precipitated by a growing foreign economy demanding more energy pushes up the domestic country's exports. Those higher exports dampen the fall in domestic output caused by higher energy prices and put more upward pressure on domestic prices.

[34] Empirical studies have produced conflicting results on whether U.S. monetary policy enhanced output's response to oil price shocks in the 1970s. See Bernanke, Gertler, and Watson (1997), Hamilton and Herrera (2004), and Kilian and Lewis (2011).

In contrast, a decrease in foreign energy supply pushes down output worldwide, which reduces international trade. If the decreases in domestic exports and domestic imports are offsetting, then an energy price increase caused by a decline in foreign energy supply has no additional effects on our model.[35]

The findings of Balke and Brown (2018) suggest our model is best at explaining the effects of an energy price increase caused by a decline in foreign energy supply. The authors' results, however, have implications for how our model's results would change when an energy price increase is caused by an increase in foreign demand for energy. Specifically, an increase in foreign demand would lead to a smaller decline in domestic output and a larger rise in the price level than a decrease in foreign supply. Those changes have implications for when the economy hits the downward rigid nominal wage constraint. The smaller decline in output means the real wage will not decrease as much in the early periods, while the larger rise in the price level implies the inflation rate initially will be higher. That combination of responses signifies the economy is less likely to hit the downward nominal wage constraint. Therefore, an energy price shock caused by a foreign supply disruption will have a larger asymmetric effect on output than an energy shock caused by an increase in foreign demand.

6. Conclusion

This paper introduces downward rigid nominal wages into a standard New Keynesian model in which energy is both a factor of production and a consumption good. An energy price increase that causes the downward nominal wage constraint to bind limits the real wage rate's decline, which forces firms to reduce output more than without the constraint. That downward constraint, however, has no meaningful impact on the real wage after an energy price decrease, so

[35]Another possibility is that an oil price increase is caused by an increase in domestic demand. For example, suppose a positive domestic or worldwide aggregate demand shock raises the demand for energy, which leads to a jump in the energy price. That higher price would raise the real marginal cost leading to a smaller increase in output and a larger rise in inflation.

output does not rise as aggressively. Thus, downward rigid nominal wages cause energy price shocks to have asymmetric effects on the macroeconomy. Our results show the degree of asymmetry depends on the labor supply elasticity, the amount of price stickiness, the steady-state inflation rate, and the degree of downward nominal wage rigidity.

The model with downward rigid nominal wages provides a theoretical explanation for the economy's asymmetric response to oil price shocks. We contend that large oil price shocks, which push the price of oil to new highs relative to recent experience, are much more likely to cause the downward nominal wage constraint to bind. For example, the 64 percent increase in the price of oil from February 1980 to February 1981 was so large that most energy-intensive firms were unable to lower wages enough to offset the jumps in their marginal costs. As such, those firms were forced to reduce their output further. The example shows that a downward rigid nominal wage constraint is a reasonable mechanism to include in any theoretical model seeking to explain output's asymmetric response to a large oil price shock.

One potential concern with our specification of downward nominal wage rigidity is that the constraint is absolute. Nominal wages are perfectly flexible, but they cannot decline below a certain level. In the real world, nominal wages likely face asymmetric adjustment costs that increase in size as nominal wages fall further below a certain threshold. That modification to our New Keynesian model would change our quantitative results, but it would not change our qualitative results. A more accurate specification of the downward nominal wage constraint is necessary when addressing questions, such as the optimal policy response to oil price shocks, where precise quantitative results matter. Those topics, however, are beyond the scope of this paper and are left for future research.

References

Aastveit, K. 2014. "Oil Price Shocks in a Data-Rich Environment." *Energy Economics* 45 (September): 268–79.

Abbritti, M., and S. Fahr. 2013. "Downward Wage Rigidity and Business Cycle Asymmetries." *Journal of Monetary Economics* 60 (7): 871–86.

Abo-Zaid, S. 2013. "Optimal Monetary Policy and Downward Nominal Wage Rigidity in Frictional Labor Markets." *Journal of Economic Dynamics and Control* 37 (1): 345–64.

Balke, N. S., and S. P.A. Brown. 2018. "Oil Supply Shocks and the U.S. Economy: An Estimated DSGE Model." *Energy Policy* 116 (May): 357–72.

Baqaee, D. R. 2020. "Asymmetric Inflation Expectations, Downward Rigidity of Wages, and Asymmetric Business Cycles." *Journal of Monetary Economics* 114 (October): 174–93.

Barattieri, A., S. Basu, and P. Gottschalk. 2014. "Some Evidence on the Importance of Sticky Wages." *American Economic Journal: Macroeconomics* 6 (1): 70–101.

Benigno, P., and L. A. Ricci. 2011. "The Inflation-Output Trade-Off with Downward Wage Rigidities." *American Economic Review* 101 (4): 1436–66.

Bernanke, B. S. 1983. "Irreversibility, Uncertainty, and Cyclical Investments." *Quarterly Journal of Economics* 97 (1): 85–106.

Bernanke, B. S., M. Gertler, and M. W. Watson. 1997. "Systematic Monetary Policy and the Effects of Oil Price Shocks." *Brookings Papers on Economic Activity* 1: 91–157.

Blanchard, O., and J. Gali. 2010. "The Macroeconomic Effects of Oil Price Shocks: Why are the 2000s so Different from the 1970s?" In *International Dimensions of Monetary Policy,* ed. J. Gali and M. Gertler, 373–428. Chicago: University of Chicago Press.

Blanchard, O., and M. Riggi. 2013. "Why Are the 2000s so Different from the 1970s? A Structural Interpretation of Changes in the Macroeconomic Effects of Oil Prices." *Journal of the European Economic Association* 11 (5): 1032–52.

Bodenstein, M., C. J. Erceg, and L. Guerrieri. 2008. "Optimal Monetary Policy with Distinct Core and Headline Inflation Rates." *Journal of Monetary Economics* 55 (Supplement): S18–S33.

Bodenstein, M., L. Guerrieri, and C. J. Gust. 2013. "Oil Shocks and the Zero Bound on Nominal Interest Rates." *Journal of International Money and Finance* 32 (February): 941–67.

Calvo, G. A. 1983. "Staggered Prices in a Utility Maximizing Framework." *Journal of Monetary Economics* 12 (3): 383–98.

Christiano, L. J., and M. Eichenbaum. 1992. "Current Real-Business Cycle Theories and Aggregate Labor-Market Fluctuations." *American Economic Review* 82 (3): 430–50.

Christiano, L. J., M. Eichenbaum, and C. L. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113 (1): 1–45.

Daly, M. C., and B. Hobijn. 2014. "Downward Nominal Wage Rigidities Bend the Phillips Curve." *Journal of Money, Credit and Banking* 46 (S2): 51–93.

Dhawan, R., and K. Jeske. 2008. "Energy Price Shocks and the Macroeconomy: The Role of Consumer Durables." *Journal of Money, Credit and Banking* 40 (7): 1357–77.

Dhawan, R., K. Jeske, and P. Silos. 2010. "Productivity, Energy Prices, and the Great Moderation: A New Link." *Review of Economic Dynamics* 13 (3): 715–24.

Dixit, A., and J. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review* 67 (3): 297–308.

Drechsel, T., and S. Tenreyro. 2018. "Commodity Booms and Busts in Emerging Economies." *Journal of International Economics* 112 (1): 200–218.

Gavin, W. T., B. D. Keen, and F. T. Kydland. 2015. "Monetary Policy, the Tax Code, and the Real Effects of Energy Shocks." *Review of Economic Dynamics* 18 (3): 694–707.

Gottschalk, P. 2005. "Downward Nominal-Wage Flexibility: Real or Measurement Error?" *Review of Economics and Statistics* 87 (3): 556–68.

Greenwood, J., Z. Hercowitz, and G. W. Huffman. 1988. "Investment, Capacity Utilization, and the Real Business Cycle." *American Economic Review* 78 (3): 402–17.

Hamilton, J. D. 1988. "A Neoclassical Model of Unemployment and the Business Cycle." *Journal of Political Economy* 96 (3): 593–617.

———. 1996. "This Is What Happened to the Oil Price/Macroeconomy Relation." *Journal of Monetary Economics* 38 (2): 215–20.

———. 2003. "What is an Oil Price Shock?" *Journal of Econometrics* 113 (2): 363–98.

———. 2008. "Oil and the Macroeconomy." *New Palgrave Dictionary of Economics,* ed. S. Durlauf and L. Blume. Palgrave Macmillan.

———. 2011. "Nonlinearities and the Macroeconomic Effects of Oil Prices." *Macroeconomic Dynamics* 15 (S3): 364–78.

Hamilton, J. D., and A. M. Herrera. 2004. "Oil Shocks and Aggregate Economic Behavior: The Role of Monetary Policy." *Journal of Money, Credit and Banking* 36 (2): 265–86.

Hazell, J., and B. Taska. 2018. "Posted Wage Rigidity." Manuscript.

Heathcote, J., K. Storesletten, and G. L. Violante. 2014. "Consumption and Labor Supply with Partial Insurance: An Analytical Framework." *American Economic Review* 104 (7): 2075–2126.

Herrera, A. M., L. Lagalo, and T. Wada. 2011. "Oil Price Shocks and Industrial Production: Is the Relationship Linear?" *Macroeconomic Dynamics* 15 (S3): 472–97.

Holden, T., and M. Paetz. 2012. "Efficient Simulation of DSGE Models with Inequality Constraints." Manuscript.

Holland, A. S. 1988. "The Changing Responsiveness of Wages to Price-Level Shocks: Explicit and Implicit Indexation." *Economic Inquiry* 26 (2): 265–79.

———. 1995. "Inflation and Wage Indexation in the Postwar United States." *Review of Economics and Statistics* 77 (1): 172–76.

Jo, Y. J. 2021. "Establishing Downward Nominal Wage Rigidity Through Cyclical Changes in the Wage Distribution." Manuscript.

Jorda, O. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1): 161–82.

Keen, B. D., and E. F. Koenig. 2018. "How Robust Are Popular Models of Nominal Frictions?" *Journal of Money, Credit and Banking* 50 (6): 1299–1342.

Kilian, L. 2008. "The Economic Effects of Energy Price Shocks." *Journal of Economic Literature* 46 (4): 871–909.

———. 2009. "Not All Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market." *American Economic Review* 99 (3): 1053–69.

Kilian, L., and L. Lewis. 2011. "Does the Fed Respond to Oil Price Shocks?" *Economic Journal* 121 (555): 1047–72.

Kilian, L., and C. Vega. 2011. "Do Energy Prices Respond to U.S. Macroeconomic News? A Test of the Hypothesis of Predetermined Energy Prices." *Review of Economics and Statistics* 93 (2): 660–71.

Kilian, L., and R. J. Vigfusson. 2011. "Are the Responses of the U.S. Economy Asymmetric in Energy Price Increases and Decreases?" *Quantitative Economics* 2 (3): 419–53.

———. 2013. "Do Oil Prices Help Forecast U.S. Real GDP? The Role of Nonlinearities and Asymmetrics." *Journal of Business and Economic Statistics* 31 (1): 78–93.

Kim, J., and F. J. Ruge-Murcia. 2009. "How Much Inflation Is Necessary to Grease the Wheels?" *Journal of Monetary Economics* 56 (3): 365–77.

———. 2011. "Monetary Policy When Wages Are Downwardly Rigid: Friedman Meets Tobin." *Journal of Economic Dynamics and Control* 35 (12): 2064–77.

Mehra, Y. P. 2002. "The Taylor Principle, Interest Rate Smoothing and Fed Policy in the 1970s and 1980s." Working Paper No. 02-03, Federal Reserve Bank of Richmond.

Mork, K. 1989. "Oil and the Macroeconomy When Prices Go Up and Down: An Extension of Hamilton's Results." *Journal of Political Economy* 91 (3): 740–44.

Plante, M. 2014. "How Should Monetary Policy Respond to Changes in the Relative Price of Oil? Considering Supply and Demand Shocks." *Journal of Economic Dynamics and Control* 44 (July): 1–19.

Ramey, V. A., and S. Zubairy. 2018. "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data." *Journal of Political Economy* 126 (2): 850–901.

Schmitt-Grohe, S., and M. Uribe. 2016. "Downward Nominal Wage Rigidity, Currency Pegs, and Involuntary Unemployment." *Journal of Political Economy* 124 (5): 1466–1514.

Sims, C. 2002. "Solving Linear Rational Expectations Models." *Computational Economics* 20 (1–2): 1–20.

Taylor, J. B. 1993. "Discretion Versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy* 39 (1): 195–214.

Wei, C. 2003. "Energy, the Stock Market, and the Putty-Clay-Investment Model." *American Economic Review* 93 (1): 311–23.

A New Measure of Central Bank Transparency and Implications for the Effectiveness of Monetary Policy*

Miguel Acosta
Federal Reserve Board

Transparency has been posited as a channel through which monetary policy is made more effective. However, empirical studies of this question and other questions concerning the role of transparency have lacked access to a time-varying high-frequency measure of transparency. This paper presents a new measure of the transparency of Federal Reserve deliberations, derived from the documents that the Fed uses to record and summarize each of its meetings. The measure—the similarity of the minutes and transcripts of each Federal Open Market Committee (FOMC) meeting—is largely, though not entirely, shaped by FOMC leadership. Monetary policy shocks have about a 40 percent larger effect on nominal and real

*This paper previously circulated under the title "FOMC Responses to Calls for Transparency" (Acosta 2015). That incarnation of the paper benefited greatly from discussions with my undergraduate honors thesis advisor, John Taylor; the support of my parents and sister: José Acosta, Mary Grebenc, and Lourdes Acosta; and the helpful comments of Lydia Cox, Geoffrey Cox, Marcelo Clerici-Arias, Chris Gust, José Gutierrez, Donna Hunter, Stephanie Kalfayan, Mark Lucianovic, Ellen Meade, the late Allan Meltzer, Elmar Mertens, Kurt von Tish, the Bing Honors College (2013) participants, and seminar participants at the Federal Reserve Board. This version was further improved by participants in the Monetary Economics Colloquium at Columbia, as well as by discussions with Pierpaolo Benigno (editor), Lydia Cox, Juan Herreño, Jennifer La'O, Frederic Mishkin, José Montiel Olea, Emi Nakamura, Andrea Prat, Ricardo Reis, Martin Rotemberg, Joseph Saia, Jón Steinsson, and several anonymous referees. A huge thanks goes to Stephen Hansen and Lucas Husted, who graciously provided me with data, and Blaine Helleloid, who provided outstanding research assistance. This work was supported financially by the Alfred P. Sloan Foundation Pre-Doctoral Fellowship in Behavioral Macroeconomics awarded through the NBER. The views expressed here are mine only, and should not be interpreted as the views of the Board of Governors of the Federal Reserve System, or of any other person associated with the Federal Reserve System. Author e-mail: miguel.acosta@frb.gov.

interest rates when the prevailing level of transparency is high, suggesting an important role for transparency in determining the efficacy of monetary policy. These effects are primarily driven by transparency about monetary policy strategies conditional on the state of the economy.

JEL Codes: E58, H83, D78, D82.

1. Introduction

Over the years, the world's major central banks have become more transparent in many respects—explicit inflation targets are well established, policy actions are announced, and forecasts are provided to the public. An aspect of transparency that is more difficult to quantify, however, is the extent to which the publicly provided rationales for policies reflect the reaction function of the monetary policy committee. As stressed by Woodford (2001, 2005) and numerous speeches by Federal Reserve Chairs,[1] communicating these rationales helps the public to better understand how policymakers might react to different future states of the economy, which in turn gives monetary policy greater control over longer-term interest rates. In this paper, I provide the first high-frequency measure of this "procedural transparency." With a quantitative measure in hand, we can then tackle the question of whether transparency makes monetary policy more effective. The second part of this paper provides an answer in the affirmative.

I construct the measure of procedural transparency—henceforth referred to as "transparency" unless otherwise noted—using the documents that the Federal Open Market Committee (FOMC) releases to the public as records of its policymaking meetings. Since 1976, the FOMC has consistently recorded nearly verbatim records of its meetings in documents called the *transcripts* and shorter summaries called the *minutes*. The transcripts, while detailed accounts of how the Committee comes to decisions, are not released until at least

[1]See, for example, Bernanke (2010, 2013) and Yellen (2013). The aspect of transparency described by these papers is not *exactly* "procedural transparency" as used here, but it is argued below that procedural transparency is the relevant form of transparency for explaining the rationale of policy, which is what these papers and speeches broadly consider.

five years after a meeting has taken place. (Before 1993, they were not expected to be released at all.) The lag between meetings and the release of the transcripts renders much of the information they contain stale in terms of understanding the Committee's current thinking. The minutes, on the other hand, are released with a much shorter lag—today, three weeks. My measure of transparency takes advantage of this timing: It is the similarity of the minutes and transcripts of each meeting, computed using natural language processing (NLP) techniques.

Minute-transcript similarity and the distribution of the topics in the minutes and transcripts are most-strongly associated with FOMC leadership. Minute-transcript similarity was at its highest during the late Greenspan years (early 2000s), and increased noticeably after 1993, the year in which the FOMC began publishing the records of its meetings. This is consistent with predictions from earlier studies of this event. The distribution of topics in the transcripts also changed significantly in 1993. Minute-transcript similarity is not predictable by macroeconomic variables, though changes in the distributions of the minutes and transcripts are. Higher transparency is also weakly associated with monetary shocks—as measured by several authors—that are smaller in magnitude, lending credence to the notion that the measure helps to inform the public about the Federal Reserve's policy. In Sections 2 and 3, I describe the construction of the measure and its properties, respectively, to establish that it provides a meaningful measure of Fed transparency.

In its most literal interpretation, minute-transcript similarity captures the overlap between the distributions of topics discussed in the transcripts and those discussed in the minutes. My measure should therefore be interpreted as a way to understand the divergence between what receives the Committee's attention when a decision is being made and how that thinking is described.[2] Given a

[2]Note that a clear, concise, and informative summary of a long discussion—a discussion possibly filled with tangents, misunderstandings, and other banter—should not be expected to cover the same topics in the same proportion, though my proposed measure would penalize such deviations. Variation in minute-transcript similarity could arise from the noise present in natural language/conversation, variation in how difficult a meeting is to describe (say, for example, because of a complicated policymaking environment), or strategic considerations regarding the transmission of information. That said, to my

higher level of transparency, then, the public and financial market participants should be able to better predict the Fed's policies, as elevated transparency implies that the Fed's communications are providing clearer insights into policymakers' thinking. This motivates the main empirical question of the paper: How does transparency affect the effectiveness of monetary policy, as measured by the pass-through from short-term nominal rates to long-term nominal *and* real rates?

Section 4 turns to the role that transparency plays in determining the effectiveness of monetary policy. Specifically, I show that the monetary policy shocks of Nakamura and Steinsson (2018) have larger effects on real interest rates when transparency is elevated. In Sections 5.1 and 5.2, I show that these results are primarily driven by transparency about monetary policy strategies conditional on the economic outlook. Additionally, the effects of monetary policy estimated by Nakamura and Steinsson (2018) are shown to be slightly downwardly biased. This arises because some of the larger monetary shocks have been delivered at times when transparency is low. But it is precisely when transparency is low that monetary policy shocks have smaller effects on real interest rates, possibly because the public cannot make as much sense of the short-run surprises as they relate to the path of future rates. These findings are robust to concerns that transparency may be proxying for some other variable—forward guidance, the state of the economy, and public uncertainty about monetary policy.

2. Measuring Procedural Transparency

2.1 Previous Measurements

The literature on central bank transparency and communication started in earnest at the turn of the century—Blinder et al. (2008) provide a thorough survey of this literature through 2008. This paper contributes to a branch of this literature concerning the

knowledge, there is no evidence that the minutes are anything but a forthright effort to summarize the transcripts (though, such evidence might be hard to find). Additionally, I find in Section 5.2 that my results are robust to using measures of transparency about obviously meaningful topics—i.e., measures that should be less affected by these concerns.

measurement of transparency—most recently treated by Dincer and Eichengreen (2014), preceded by Eijffinger and Geraats (2006),[3] who based their measures on the Geraats (2001, 2002) definitions of different aspects of central bank transparency. This paper centers around a particular component of Geraats' *procedural transparency* that concerns central bank accounts of deliberations, the measurement of which has, thus far, focused primarily on fairly aggregated and slow-moving measures of the timeliness and informativeness of central bank communications.[4] For example, the relevant measure of procedural transparency in Dincer and Eichengreen (2014) is a binary indicator of whether "the central bank give[s] a comprehensive account of policy deliberations (or explanations in case of a single central banker) within a reasonable amount of time." In this paper, the use of natural language processing techniques allows these shortfalls to be circumvented by using the text of each FOMC meeting. Because the measure is constructed from text, it reflects fairly detailed changes in communications and transparency (it is continuous, not discrete). Because it changes at every FOMC meeting, it gives a high-frequency measure of transparency that, in practice, changes much more often than previous measures that primarily capture large regime changes in communications policies.

2.2 How to Measure Transparency

My proposed measure of procedural transparency is the similarity between the minutes and transcripts of each FOMC meeting, which I will refer to as "minute-transcript similarity" or simply "transparency" for the rest of the paper. The transcripts contain a nearly verbatim record of each FOMC meeting, yet are not released until at least five years after an FOMC meeting has taken place—before 1993, they were not released at all.[5] The minutes are

[3]Dincer and Eichengreen (2014) also thoroughly discuss previous literature.

[4]See Appendix B for a discussion about the different types of transparency relevant for a central bank. Procedural transparency, as defined by Geraats (2002), is the description of how monetary policy decisions are made, which is achieved in part through the publication of records of the deliberative process.

[5]"Nearly verbatim" refers to the fact that the exact words are lightly edited. From the Federal Reserve Board's website: *The most detailed record of FOMC*

shorter summaries—typically 10–20 pages to summarize 100–200 page transcripts—released three weeks after each meeting has taken place.[6] Evidence suggests that the minutes are intended to be accurate portrayals of what was discussed at each meeting, and are not intended to obfuscate the content of the discussion. Chair Yellen, questioned about this in her June 2016 press conference, responded that "the minutes are always—have to be—an accurate discussion of what happened at the meeting."[7] In addition, the FOMC has to vote on the minutes, presumably reducing the possibility of systematic obfuscation.

The first step in computing minute-transcript similarity is to represent each transcript and minutes as a distribution over a finite number of topics, using Latent Dirichlet Allocation (LDA). The similarity between the minutes and the transcripts for a particular meeting is the Kullback-Leibler similarity of the distributions of the two documents—a measure that lies in the interval $[0, 1]$. The measure bears a striking resemblance to Chair Greenspan's interpretation of how the public understood FOMC communications, which he voiced in the September 2003 FOMC meeting: "A number of those in the market don't listen to the subtleties; they just take note of how much time we are spending talking about a particular subject" (FOMC Transcripts 1976–2008).

meeting proceedings is the transcript. Beginning with the 1994 meetings, the FOMC Secretariat has produced the transcripts shortly after each meeting from an audio recording of the proceedings, lightly editing the speakers' original words, where necessary, to facilitate the reader's understanding. Meeting participants are given an opportunity within the subsequent several weeks to review the transcript for accuracy.

For the meetings before 1994, the transcripts were produced from the original, raw transcripts in the FOMC Secretariat's files. These records have also been lightly edited by the Secretariat to facilitate the reader's understanding. In addition, where one or more words were missed or garbled in the transcription, the notation "unintelligible" has been inserted. In some instances, words have been added in brackets to complete a speaker's apparent thought or to correct an obvious transcription error or misstatement.

[6] The release lag has changed over the years—see Appendix A for more details.

[7] This is not a recent sentiment; in 2003, the Federal Reserve Bank of Kansas City president stated in an FOMC meeting that "the minutes are our representation of the discussions that occurred at the meeting" (FOMC Transcripts 1976–2008, Sept. 2003).

Before detailing the procedure used to compute minute-transcript similarity, it is important to understand what this measure is and its potential shortcomings. The measure represents the extent to which the content of the transcripts is reflected proportionately in the minutes.[8] Because LDA represents each document as a distribution over topics, the mass placed on each topic describes the amount of the document devoted to that topic. Thus, only when the minutes devote exactly the same amount of space to each topic as the transcripts will the measure equal unity. The view taken in this paper is that it is not the job of the minutes writers to editorialize the FOMC's discussion—e.g., to eliminate the side of a debate that does not ultimately "win"—but instead to convey the discussion accurately. Put differently, minutes that fully communicate FOMC discussions are taken as transparent.

That said, the inclusion of obviously irrelevant discussions—e.g., "when should we break for lunch"—should not be a necessary condition for minutes to be transparent. Of course some divergence between the two documents should be expected—while a conversation might be centered around a topic, the actual words used or topics discussed might only noisily represent that topic—owing to, for example, digressions or misunderstandings. The underlying assumption I make is that this noise is fairly constant over time, only affecting the level of my measure and not its changes. In Section 5.2, I show that my empirical findings are robust to more-narrowly defined measures of transparency that are constructed using economically important topics (and should thus be relatively free from this type of noise). This is a benefit of using the fairly complex language model described in the next section—it allows for documents to be analyzed at the topic level and thus separate out the types of discussions that can add noise to my measure.

2.3 Language Model

Several steps are involved in computing the topic distribution for each FOMC transcript and minutes, with the ultimate goal being a representation of a document into well-understood topics

[8]Note that the measure does not consider the similarity of documents *at different meetings*, only the similarity of documents related to the same meeting.

in a high-frequency, holistic, and interpretable way. A drawback of the approach used here—along with the vast majority of NLP techniques—is that documents are represented as "bags of words," i.e., the order of words does not matter. The only features of a document that are retained are counts of the number of times that each unique word appears in that document. This is a necessary evil when working with a large body of text—the 600 documents used in the analysis here contain 13 million individual, and 170,000 unique, words.[9]

After converting the documents into computer readable formats, these raw text files are "preprocessed" using several techniques that are standard in working with natural language. This preprocessing achieves three goals. The first goal is to reduce the effect of errors that might arise from working with a data source that may contain typographical errors or other errors arising from the fact that several of the documents had to be converted from typewritten documents. To that end, only the letters of the alphabet are retained, and every unique word must appear at least three times over the entire corpus, otherwise it is dropped. Words shorter than 3 characters or longer than 15 are also dropped—the intention of the latter being to remove words that may have been accidentally concatenated.

The second goal of the preprocessing is to reduce the noise that arises from grammatical constraints: The words *increase, Increase!, increasing,* and *increased* all convey essentially the same meaning, yet a simple numerical representation of the words in the document might treat them as completely different words, since it knows no better. To that end, words are stemmed to their lexical root, so that in the example above, all occurrences of the "increase" words are stemmed to *increas.*[10] Terms in a "stoplist" are also excluded. As is customary, this list contains common words that contribute little meaning to the documents, since they are used so often. The excluded words are the "generic," "dates and numbers," and "geographic" lists from Loughran and McDonald (2011), who carefully constructed these lists to be relevant in a context of finance.

[9]This is based off of simply splitting the documents by whitespace, i.e., with no preprocessing.

[10]Stemming is performed using the Lancaster Stemmer as implemented in Python's Natural Language Toolkit.

The final goal of preprocessing is to reduce the noise that arises when ideas need to be mapped into words, and vice versa. This goal is addressed via an application of Latent Dirichlet Allocation, developed by Blei, Ng, and Jordan (2003).[11] Because LDA enjoys widespread use in the NLP community, and even within the economics literature, the treatment here is brief. First, the observed corpus contains words, with $w_{d,n}$ being the n^{th} word in the d^{th} document ($d \in \{1 \ldots, D\}$), where each document has N_d words (so that $n \in \{1, \ldots, N_d\}$, $\forall d$). This is all that is observed. LDA posits each document as a distribution over a fixed number, K, of topics—K is chosen by the researcher. Topics are in turn distributions over the V unique terms in the corpus. More precisely, each document δ_d is a draw from a Dirichlet distribution, a distribution over vectors that lie in the K-simplex. The distribution, δ_d, that is drawn from the Dirichlet is a latent variable. The same is true of topics: Each topic, ϕ_k, is a draw from a V-dimensional Dirichlet. With a topic distribution in hand, each observed document (which has a fixed length, N_d) is populated one word at a time. For the n^{th} slot of document d, a topic is drawn from a multinomial distribution, with parameter δ_d. Thus, if δ_d is heavily concentrated on topic 1, then several words will be drawn from topic 1. The drawn topic, $z_{d,n} \in \{1, \ldots, K\}$, is then used to draw a word from a multinomial distribution with parameter $\phi_{z_{d,n}}$. So, if a topic has a distribution that places heavy weight on "whale," then "whale" will come up often when that topic is drawn, and will thus show up often in documents that have a high probability placed on that topic. The implementation here—including the choice of the priors for the Dirichlet distribution—follows very closely that of Hansen, McMahon, and Prat (2017), who estimate LDA on a subset of the FOMC transcripts that I consider using a Gibbs sampler outlined in Griffiths and Steyvers (2004).[12] The

[11] At issue here are the problems of synonymy and polysemy. Polysemy occurs when one word can describe many concepts. For example, polysemy would lead the documents [I read a book.] and [I'll book a hotel.] to look more similar than an English-speaking human might think. Next, synonymy occurs when any concept can be expressed using many different words. The document [I'll make lodging arrangements.] and the hotel document from above would look dissimilar, despite conveying the same idea.

[12] And very graciously provided by the authors at https://github.com/sekhansen/text-mining-tutorial.

number of topics, K, is set to 50, chosen using a fivefold cross-validation technique similar to that outlined in Hansen, McMahon, and Prat (2017) and described in Appendix C.

Rather than estimating the topic model over the complete documents in my corpus, I instead begin by splitting these documents (the minutes and transcripts) into sentences using a grammatical sentence parser.[13] I then estimate the word and topic distributions over every sentence in the FOMC minutes and transcripts between 1976 and 2014. In the notation established above, then, d indexes sentences so that each sentence has an estimated distribution over topics. With these estimated distributions, I then estimate the topic distributions for each transcript as a whole ($\theta_t \in \mathbb{R}^K$), and each of the minutes as a whole ($\mu_t \in \mathbb{R}^K$). This follows the approach of Hansen, McMahon, and Prat (2017), the purpose being to have each document focused on a small number of topics, in hopes that the latent topics can be determined more easily.[14]

The "Top 10 Words" column of Table 1 contains the 10 words that most prominently contribute to the composition of each topic. Formally, these are the words corresponding to the 10 largest elements of ϕ_k for each topic k. For purposes of interpretation I show the estimated topics for a smaller model with $K = 30$ (my empirical results in Section 4 are nearly unchanged quantitatively using the larger model) that I estimated over the sample beginning in 1995, since that forms the basis of my empirical estimates in Section 4. In general the topics seem intuitive—the first four topics might be subjectively called topics about labor, credit markets, housing markets, and policy statement language. Not every topic is directly related to an economic concept—the fifth topic contains words that might be used in a debate regarding policy communications.

[13] I parse the minutes and transcripts into sentences using the English probabilistic context-free grammar developed by Klein and Manning (2003) as implemented in the Stanford Parser Java package. This sentence parser uses rules of English grammar to split sentences, as opposed to simple rules based solely on punctuation. This, for example, avoids erroneously splitting sentences at decimal points or after abbreviations (e.g., "Ms.").

[14] This substantially increases D to about 260,000, though there are 300 minutes and 300 transcripts. This approach of using already-estimated topics to estimate the topic distribution of an excluded document underlies the right panel of Figure B.1, with the excluded document reading "transcripts minutes record policy actions memorandum discussion communications."

Table 1. Topic Descriptions

Topic	Top 10 Words										Dissent	FFR	GDP	Inflation
1	level	unemploy	percent	rate	employ	low	averag	high	declin	remain				
2	credit	bank	spread	loan	debt	investor	mortgag	bond	larg	money			√	
3	hous	sale	sector	inventori	activ	declin	construct	home	start	weak			√	
4	meet	statement	altern	committe	chang	languag	fomc	paragraph	direct	word	√		√	
5	think	make	commun	public	view	way	help	decis	use	forward	√			
6	market	financi	labor	particip	improv	condit	fore	gener	substanti	note	√			
7	purchas	asset	treasuri	end	secur	program	hold	back	reduc	size			√	
8	spend	busi	consum	invest	incom	capit	household	confid	increas	continu	√			
9	risk	balanc	concern	view	side	possibl	sheet	downsid	outlook	signific				
10	time	take	need	think	begin	adjust	process	place	step	normal				
11	chang	littl	move	think	bit	probabl	reason	case	quit	fact		√	√	
12	like	look	see	come	go	time	forward	back	number	reason				
13	polici	monetari	action	fiscal	accommod	tighten	eas	appropri	addit	support	√	√	√	
14	line	dollar	right	panel	lrb	rrb	shown	foreign	show	left		√		
15	term	long	expect	run	longer	short	consist	near	time	object	√			
16	growth	continu	economi	pace	slow	moder	product	expans	trend	econom		√	√	
17	reserv	bank	oper	feder	system	account	open	central	facil	direct	√		√	
18	recent	data	month	suggest	year	past	indic	survey	seen	evid				
19	peopl	lot	get	work	think	way	talk	just	say	thing	√			
20	price	increas	cost	higher	pressur	rise	energi	oil	wage	declin		√	√	√
21	economi	uncertainti	problem	situat	world	face	shock	deal	event	creat				
22	forecast	project	growth	real	gdp	greenbook	staff	point	half	revis	√			
23	effect	demand	import	factor	reflect	part	posit	suppli	export	state		√	√	
24	inflat	expect	percent	measur	core	remain	pce	rang	higher	low				√
25	rate	fund	point	interest	basi	feder	target	lower	fed	rais	√			
26	report	good	product	industri	firm	district	manufactur	order	contact	nation	√			
27	differ	use	model	base	output	gap	rule	structur	estim	actual				
28	question	presid	issu	chairman	discuss	comment	said	governor	mention	ye			√	
29	think	go	want	say	know	just	dont	get	realli	way			√	
30	econom	period	committe	outlook	member	develop	stabil	inform	intermeet	consider				√

To better aid in categorizing the estimated topics, the last four columns of Table 1 indicate which of the topics are useful for "predicting" a few external observable variables, denoted by e_t. Specifically, I gather the number of dissenting votes at each FOMC meeting from Thornton and Wheelock (2014), and the change in the target of the federal funds rate (FFR), GDP growth, and inflation from the Federal Reserve Economic Data (FRED) data portal.[15] I then estimate which topics are useful predictors of these variables by estimating the following regression using the Lasso objective function:

$$e_t = \beta_0 \mathbf{1}\{\text{ELB}\}_t + \sum_{k=1}^{K} \beta_k \delta_t^k + \text{error}_t,$$

where δ_t^k is the estimated presence of each topic k in document t. I estimate this regression twice: once using the presence of topics in the minutes (i.e., replacing δ_t^k with μ_t^k) and once using their presence in the transcripts (i.e., replacing δ_t^k with θ_t^k). I take the union of the selected topics between these two regressions. I select the Lasso regularization parameter using tenfold cross-validation.[16] This again helps to shed light on the estimated topics. The selected topics for inflation are the most intuitive, with words like "price, inflat, econom" being the top words for the selected topics (20, 24, and 30). Instead, topics about appropriate communication and policy (4, 5, 25) and the general outlook for the economy (most of the other selected topics) can predict the number of dissents at each meeting.

2.4 Transparency Index Definition

Ultimately, the object of interest in this paper is not the topics themselves, but rather the relative entropy of the minutes for the

[15]In terms of FRED mnemonics, the FFR target is `DFEDTAR` when it is available and the midpoint of `DFEDTARL` and `DFEDTARU` when it is not; GDP growth is the four-quarter difference in the log of `GDPC1`; and inflation is the 12-month difference in the log of `PCEPI`.

[16]I allow the intercept to vary based on whether the federal funds target is at its effective lower bound (ELB), since the FFR is one of my target variables. I implement the estimation and cross-validation using the LassoCV module in Python's sklearn package; the sample is split into 10 disjoint subsets that are the same for each e_t.

transcripts (i.e., the Kullback-Leibler similarity) and the entropies of the minutes and transcripts on their own. Specifically, given the topic distribution of the transcripts for the FOMC meeting occurring at time t, θ_t, and the topic distribution of the minutes for the same meeting, μ_t, I define the three quantities:

$$\text{Minute-Transcript Similarity:} \qquad \tau_t \equiv \exp\left[-\sum_{k=1}^{K} \mu_t^k \ln\left(\frac{\theta_t^k}{\mu_t^k}\right)\right]$$

$$\text{Entropy of the Minutes:} \qquad H(\mu_t) \equiv -\sum_{k=1}^{K} \mu_t^k \ln\left(\mu_t^k\right)$$

$$\text{Entropy of the Transcripts:} \qquad H(\theta_t) \equiv -\sum_{k=1}^{K} \theta_t^k \ln\left(\theta_t^k\right)$$

with θ_t^k being the k^{th} element of θ_t, and analogously for μ_t. The first measure is the negative exponential of the Kullback-Leibler divergence, a distance function for distributions—intuitively it describes the information loss from assuming the truth is the minutes, when it is really the transcripts. The other two measures, the entropies of the minutes and transcripts, are the expected values of the information content of a random variable that is distributed according to μ_t and θ_t. Entropy in this case achieves a maximum when both distributions place equal mass on each topic, i.e., $\mu_t^k = \frac{1}{K}, \forall k$, and decreases as mass moves away from certain topics and concentrates on others.[17] Thus, entropy in this context can be cast intuitively as a measure of how dispersed a conversation/document is—the lower the entropy, the more concentrated the discussion.

For the construction of the procedural transparency measure, the body of documents under consideration consists of all Records of Policy Actions (ROPA, an older version of the minutes), and minutes and transcripts from meetings physically held in Washington,

[17] Information about a variable drawn from a uniform distribution is more valuable than information about a variable drawn from a point mass—one already knows where the point mass is, but has no idea where the uniform variable is.

D.C., between April 1976 and December 2014, where procedural information (voting records, attendance) is removed.[18]

3. Notable Features of the Minutes, Transcripts, and Procedural Transparency

This section presents my measure of procedural transparency and three exercises meant to better understand it. First, in Section 3.1, I compare the measure with several external variables, in order to understand its systematic components. My primary finding is that fluctuations in the measure, while correlated with some of these external variables, are not a proxy for something simpler. Next, in Section 3.2, I highlight, anecdotally, that the measure captures meaningful differences between the minutes and transcripts. Finally, in Section 3.3, I argue that my measure—though not directly observable in real time—might be roughly observable to the public. This is done using a newspaper-based measure of central bank transparency.

3.1 Correlations with External Variables

Figure 1 presents the time series of minute-transcript similarity, and Figure 2 contains the entropies of the minutes and transcripts. All three measures are rather noisily distributed around slower-moving trends, shown in the images as 12-meeting trailing moving averages. Owing to concerns—mentioned in Section 2.2—that the raw measures are likely influenced by idiosyncratic noise arising from the noise inherent in natural language, the moving averages of these series are the main measures considered for the rest of the paper.

In order to analyze the series more carefully, Tables 2 and 3 present the results of regressing the standardized measures on several variables. The first, Table 2, shows the regression of these communications variables on the other variables one at a time. The

[18]See Appendix A for a discussion of the various documents released by the FOMC since its inception. For the modern-day minutes (1993–present), all words prior to the paragraph that typically begins with "The information reviewed at the x meeting . . . " (now labeled "Staff Review of the Economic Situation") are removed in order to make these documents look like the ROPA. This also keeps the content of the minutes looking relatively similar over the year, since the first meeting of each year contains discussions of procedural matters (see Meade, Burk, and Josselyn 2015 for more information on the content of the minutes).

Figure 1. Procedural Transparency

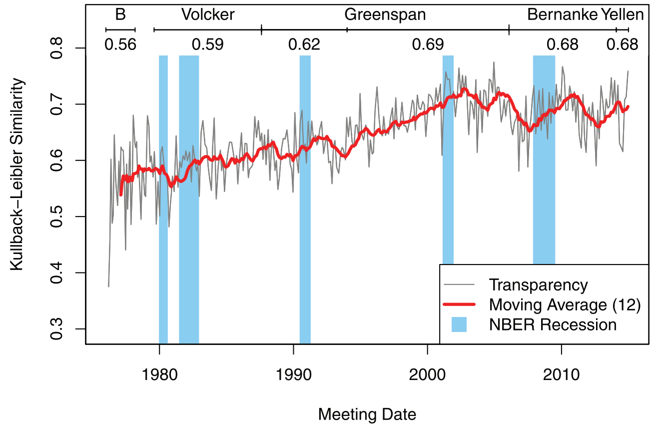
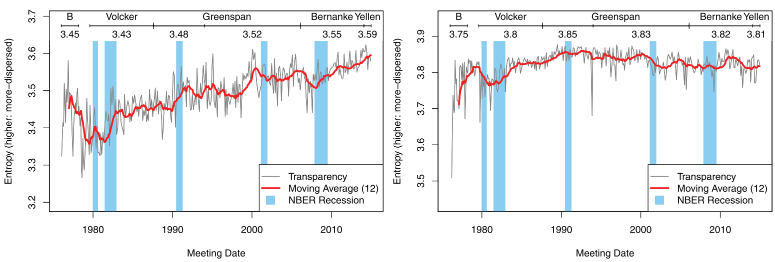


Figure 2. Entropy of the Minutes (left) and Transcripts (right)



lessons here are largely consistent with the results in Table 3, which estimates the coefficients jointly for variables that span the entire 1976–2014 sample. In Table 2, the variables are standardized so that the coefficients represent correlation coefficients; in Table 3, measures are scaled so that they can be interpreted as the number of standard deviations by which the measure moves when the variable in the row increases by one unit.[19] The regressions highlight some notable features of the series.

First, the three measures all contain positive linear trends to different degrees of statistical significance—between the two tables, the

[19] Over the sample period used for that regression.

Table 2. Regression Coefficients for Communications Variables

	Transparency		**Minutes Entropy**		**Transcripts Entropy**	
		MA		**MA**		**MA**
t	**2.50**	**3.24**	**2.63**	**3.23**	1.07	1.41
Transparency MA	**0.33**	**0.43**	0.09	0.14	0.18	0.19
Minutes Ent. MA	0.13	0.17	0.26	0.43	0.17	0.16
Transcript Ent. MA	0.04	0.09	0.08	0.07	**0.61**	**0.92**
TT State	–0.14	–0.11	–0.10	–0.07	0.12	0.30
RR Shocks_{t+1}	–0.03	–0.04	–0.03	–0.02	0.07	0.06
$\lvert\text{RR Shocks}_{t+1}\rvert$	–0.08	**–0.16**	**–0.13**	**–0.22**	**–0.13**	**–0.21**
NS Policy Shock_{t+1}	0.06	0.13	0.06	0.11	0.06	**0.22**
$\lvert\text{NS Policy Shock}_{t+1}\rvert$	0.13	–0.01	–0.05	–0.32	0.08	–0.03
FFR	**–0.55**	**–0.77**	**–0.70**	**–0.88**	–0.20	**–0.37**
Change in FFR	–0.07	–0.04	–0.09	–0.04	0.01	–0.04
|Change in FFR|	**–0.17**	**–0.23**	**–0.27**	**–0.29**	**–0.24**	**–0.28**
Δy_t	–0.20	–0.16	–0.19	–0.14	0.06	0.16
$\mathbb{E}_{GB}[\Delta y_{t+1}]$	0.02	0.13	0.16	**0.30**	0.07	0.08
Unemployment (u)	–0.23	–0.24	–0.10	–0.13	–0.20	**–0.35**
$\mathbb{E}_{GB}[u_{t+1}]$	–0.26	–0.28	–0.18	–0.24	–0.22	**–0.39**
π_t^{PCE}	**–0.61**	**–0.75**	**–0.68**	**–0.80**	**–0.41**	**–0.60**
$\mathbb{E}_{GB}[\pi_{t+1}]$	**–0.56**	**–0.70**	**–0.67**	**–0.78**	**–0.46**	**–0.57**

Note: *Dependent Variable:* This table reports coefficients of univariate regressions of the three communications measures (transparency, and the entropies of the minutes and transcripts) on the variables in the rows of the table. The regressions are shown for the moving average of each measure, and the level of the measure. In all but the first row the relevant communication measures have been linearly detrended and standardized. *Independent Variables:* The variable t is a linear trend. *TT State* is the "state of the economy" variable of Tenreyro and Thwaites (2016); "RR" are the monetary policy shocks of Romer and Romer (2004), updated through the sample; NS Policy Shock and NS Fed Funds Shock are the monetary and federal funds futures shocks of Nakamura and Steinsson (2018); FFR is the federal funds target (or the midpoint of its target range, or the actual value when neither is available); Δy_t is the annualized quarterly growth rate of real GDP from FRED; u_t is the civilian unemployment rate from FRED; π_t^{PCE} is the annualized quarterly growth rate of the PCE price level; the rows $\mathbb{E}_{GB}[x_{t+1}]$ correspond to the Greenbook forecast of x in the quarter following the FOMC meeting at time t (these correspond to the macroeconomic series from above, except that CPI inflation is used instead of PCE in order to have a longer sample from the Greenbook). The monetary policy shocks are timed such that the regression corresponds to the transparency of the minutes prevailing immediately before the shock is emitted. All variables have been standardized over the regression sample so that the coefficients reflect correlation coefficients. *Sample:* The sample sizes [for the moving averages] are as follows: 74 for the NS policy shocks (Jan. 2004–Mar. 2014) to match the results from Section 4; 276 [265] for the RR shocks (Mar. 1977–Oct. 2008); 261 for $\mathbb{E}_{GB}[x_{t+1}]$ (Oct. 1979–Dec. 2014); and 326 [315] for all other variables (Mar. 1977–Dec. 2014). **Bolded estimates** are statistically significant at at least the 5 percent level, calculated using heteroskedasticity and autocorrelation-consistent asymptotic standard errors with the automatic lag selection method of Newey and West (1994), as implemented by Baum, Schaffer, and Stillman (2010).

Table 3. Jointly Testing Correlates of Communications Variables

	Transparency		Minutes Entropy		Transcripts Entropy	
t	0.01**	(0.01)	0.01***	(0.00)	0.00	(0.01)
Burns (70–78)	2.24**	(1.06)	0.08	(0.41)	–0.78	(1.00)
Miller (78–79)	2.19**	(0.86)	–0.38	(0.32)	0.48	(1.01)
Volcker (79–87)	1.52**	(0.61)	–0.52**	(0.25)	–0.38	(0.70)
Greenspan (87–06)	1.05**	(0.42)	–0.42**	(0.20)	0.57	(0.43)
Bernanke (06–14)	0.05	(0.23)	–0.37***	(0.12)	0.48*	(0.27)
Yellen (14–18)	—	—	—	—	—	—
Broida (73–78)	—	—	—	—	—	—
Altmann (78–83)	0.07	(0.21)	–1.01***	(0.20)	0.76**	(0.37)
Axilrod (83–86)	0.03	(0.23)	0.61**	(0.25)	1.14***	(0.44)
Bernard (86–87)	0.22*	(0.12)	–0.06	(0.09)	0.30	(0.20)
Kohn (87–02)	0.13**	(0.06)	–0.12**	(0.06)	–0.11	(0.09)
Reinhart (02–07)	0.15	(0.19)	–0.30*	(0.16)	–0.97***	(0.35)
Madigan (07–10)	–0.48***	(0.16)	–0.58***	(0.13)	0.13	(0.30)
English (10–15)	–0.10	(0.19)	0.34**	(0.13)	–0.07	(0.28)
FFR	0.02	(0.04)	–0.03	(0.02)	–0.02	(0.03)
ΔFFR	–0.03	(0.04)	0.01	(0.02)	–0.04	(0.03)
$\|\Delta$FFR$\|$	–0.05	(0.04)	0.00	(0.00)	–0.01	(0.01)
Post-1993	0.33	(0.30)	–0.07	(0.23)	–0.58**	(0.27)
TT State	–0.39	(0.30)	–0.06	(0.20)	1.04***	(0.34)
Δy_t	0.13	(0.09)	–0.03	(0.02)	–0.03	(0.03)
Unemployment	–0.07	(0.07)	0.05*	(0.03)	0.09	(0.07)
π_t^{PCE}	0.01	(0.08)	0.02	(0.04)	0.08	(0.06)
Cons.	–3.00***	(1.13)	–0.46	(0.75)	–3.31*	(1.87)
Chair	0.000		0.000		0.000	
Secretary	0.001		0.000		0.000	
Fed Funds	0.632		0.050		0.577	
Macro Vars.	0.447		0.531		0.274	
R^2	0.935		0.868		0.954	
N	315		315		315	

Note: The table shows results for regressions of the three communications variables (in the columns) on dummies of the sitting chairs of the FOMC (Burns–Yellen); dummies for the secretaries (Broida–English); and several other variables defined in Table 2. Chairs and secretaries have their years in office in parentheses. The rows with italicized labels contain p-values for tests of joint significance of groups of variables in the regression: *Chair* jointly tests the chair dummies; *Secretary* tests the secretary dummies; *Fed Funds* tests the three FFR variables; and *Macro Vars.* tests output growth, unemployment, and inflation. These p-values, as well as the standard errors in parentheses, are computed using heteroskedasticity and autocorrelation-consistent asymptotic standard errors with the automatic lag selection method of Newey and West (1994), as implemented by Baum, Schaffer, and Stillman (2010). The communications variables are standardized to be mean zero with unit standard deviation over the regression sample, which is each regularly scheduled FOMC meeting between March 15, 1977 and December 13, 2014 (the meetings between April 1976 and March 1977 are dropped when the moving average is formed).

minutes and transcripts have generally increased in their breadth of coverage (higher entropy), and transparency has generally increased. One notable jump—albeit statistically insignificant—is that procedural transparency increased after the FOMC became aware that its transcripts would be released to the public in 1993, and the transcripts also became more focused (lower entropy). The fact that this event may have had an effect on the content of the transcripts—through its effect on the Committee's deliberation—is the subject of a literature that has largely concluded (with the exception of the final paper) that the event caused a move towards less debate and more formal discussions (like what might be found in the minutes); see Meade (2005), Meade and Stasavage (2008), Acosta (2015), Egesdal, Gill, and Rotemberg (2015), Hansen, McMahon, and Prat (2017), and Woolley and Gardner (2017).

Next, the sitting secretary of the FOMC is strongly correlated with the level of all three measures—the p-value for a test that these variables are jointly non-zero is negligible in all cases. This is encouraging, for it suggests that the person in charge of overseeing the creation of each document has a role in determining its properties. The sitting chair is also correlated with all three measures. The Madigan dummy—a proxy for the Great Recession—in the minutes regression shows that the minutes became much more focused during the crisis and its aftermath.

By and large, the measures are not predicted by current macroeconomic events and monetary policy once other controls are included, as can be seen in Table 3. Unconditionally the variables tend to be countercyclical: Transparency is lower when the federal funds rate and inflation are high, and both the minutes and transcripts become more focused during these times. The similarity between the minutes and the transcripts also tends to decrease in meetings preceding large (in absolute value) monetary policy shocks. In the case of the shocks of Nakamura and Steinsson (2018), this implies that markets are more surprised at time $t + 1$ when the minutes from time t were less informative about the transcripts.

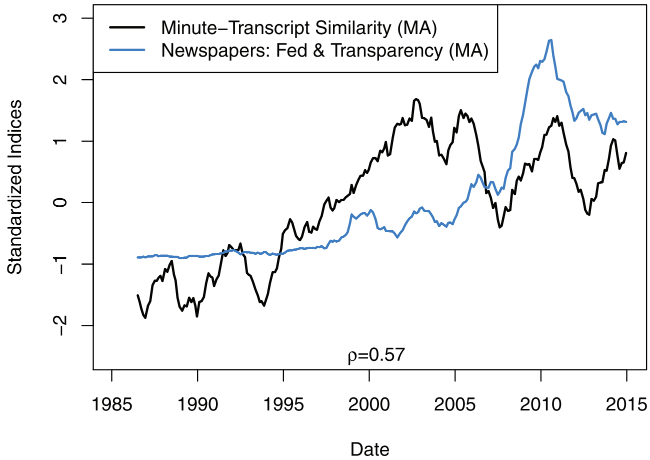
3.2 Anecdotal Understanding of the Measure

A less-systematic approach to understanding what the transparency measure is capturing is to read the underlying documents. Natural places to start are the points where the measure was particularly low and high. The lowest post-Volcker observation occurred in the December 1989 meeting, though the story starts in October 1989 with a meeting of the G-7. In that meeting, the parties involved decided that the dollar was overvalued, and agreed to a coordinated action to flood the market with dollars in order to remedy this. In the October 1989 FOMC meeting, shortly after this coordinated action began, there was a debate about whether this action interfered with the Fed's statutory mandate to achieve price stability. Some parties were concerned that the Fed would be "implicated in talking out of one side of [its] mouth about price stability goals and yet agreeing to constantly flooding the market with dollars" (FOMC Transcripts 1976–2008, Governor Johnson). Others, like President Guffey, felt differently, stating, "I'm not terribly concerned about the price stability issue in the sense that with sterilized intervention I think for some long period in the future we can go about a price stability objective without much problem" (FOMC Transcripts 1976–2008, President Guffey). At the end of the discussion, Vice Chair Corrigan suggested that the staff prepare a "presentation for the Committee where [it] would take a look at this question of price stability in five years in some systematic way." The December 1989 meeting contained that presentation and ensuing discussion, though there was no mention of it in the minutes of that meeting, this perhaps owing to President Guffey's October concern about "bringing this issue to a confrontational stage outside the confines of this Committee and the Treasury." This episode contrasts with the meeting of December 2004—the meeting with the highest level of transparency—in which the Committee undertook a lengthy discussion concerning communications policy. Specifically, they discussed the possibility of accelerating the release of the minutes from six to three weeks—a policy they subsequently implemented with the minutes of that meeting—and also made a record of this discussion in the minutes. These episodes highlight the fact that the measure captures meaningful discrepancies between the minutes and the transcripts.

3.3 Observability of the Measure

Lastly, Figure 3 shows the transparency measure alongside a measure of transparency derived from newspapers. Specifically, in the spirit

Figure 3. Monetary Policy Transparency in the News



of Baker, Bloom, and Davis (2016), Husted, Rogers, and Sun (2020) construct a measure of monetary policy uncertainty by counting the number of articles in a given period of time that appear in major newspapers containing the terms "uncertainty," "monetary policy," and "Federal Reserve."[20] The measure is divided by the number of articles that contain "Federal Reserve" for each newspaper in each period, in order to control for the volume of articles over time and the different focuses of each newspaper. After scaling the normalized counts by newspaper to have unit variance, the resulting series are summed to form the monetary policy uncertainty index. The blue line in Figure 3 is constructed in nearly the same way, with the exclusion of "monetary policy" and "uncertainty" and the inclusion of "transparent" or "transparency."[21] The resulting series is positively and significantly correlated with the transparency measure derived from the minutes and transcripts. Additionally, the moving average of the minute-transcript similarity is more-highly correlated with the newspaper-based measure (and its moving average) than the raw measure. This suggests that minute-transcript similarity, and its moving average, despite being based on a document that is not visible to the public, is something that is in some way

[20]They also include close synonyms of these three terms.
[21]A huge thanks to Lucas Husted, who constructed the new index.

observable, given its positive relationship with this clearly observable newspaper-based measure. Perhaps this arises because much of the variation in transparency is driven by the sitting chair, an easily observable variable.

4. Effectiveness of Monetary Policy

Views on the role of transparency in monetary policymaking have evolved greatly over the last 50 years. Today, transparency is often touted as a means through which monetary policy is made more effective—in 2013, then-Chair Ben Bernanke stated in a speech that "transparency about the framework of policy has aided the public in forming policy expectations, reduced uncertainty, and made policy more effective." In this section I address this question—whether transparency makes monetary policy more effective—and provide evidence suggesting that it does.

4.1 Defininig Effectiveness

The first question that arises when seeking an empirical answer to this question is how to define monetary policy effectiveness. One answer is a policy that allows a central bank to achieve its objective, such as price stability. Blinder et al. (2008) discuss the literature that has taken this approach. A general problem that arises is that establishing causal inference is challenging. Another possibility is that effective monetary policy is able to affect market expectations—Blinder et al. (2008) discuss the empirical literature that largely supports this proposition. In this vein, using a high-frequency identification strategy, Nakamura and Steinsson (2018) establish a causal link between monetary policy and real interest rates: The more surprising the monetary policy announcement, the greater the movement in real interest rates.

While these empirical studies discussed in Blinder et al. (2008), Nakamura and Steinsson (2018), and their predecessors have brought a deepened understanding of the effects of monetary policy and its communication, they say little about the role that transparency plays in determining these effects. As it pertains to transparency about the decisionmaking process, the answer is not obvious. An important characteristic of each meeting for understanding its transparency is the context in which it occurred—the state of the economy and the committee, for example. In a more-complicated policy environment, when the content of the minutes are more-heavily scrutinized,[22] high transparency might increase or decrease uncertainty and, thus, the ability of policy to have any effects. Using cross-sectional variation in transparency, Naszodi et al. (2016) highlight that transparency reduces forecast uncertainty, while increased volume of communication can have the opposite effect, as Lustenberger and Rossi (2020) find. That said, previous work has postulated that increased transparency should enhance policymakers' abilities to affect the real economy through longer-term interest rates (e.g., Woodford 2005), and the following exercises provide evidence supporting this hypothesis.

4.2 Empirical Strategy

My empirical strategy builds off of the work of Nakamura and Steinsson, who estimate the following equation:

$$\Delta i_t = a + b\,\varepsilon_t + \text{error}_t, \tag{1}$$

where i_t stands for a multitude of nominal and real interest rate forwards and yields. I take estimates of daily nominal rates from Gürkaynak, Sack, and Swanson (2007), and real rates from Gürkaynak, Sack, and Wright (2010).[23] I start my analysis in 2004 since that is when data on all interest rates under consideration are available, as discussed in Gürkaynak, Sack, and Wright (2010). The variable ε_t is a high-frequency monetary policy shock identified as the first principal component of the change in federal funds and Eurodollar futures out to four quarters.[24] This change is taken over a narrow window around FOMC statement releases. Nakamura and Steinsson find that, in response to a monetary shock, nominal and

[22] A simple Google Trends search confirms this—"Fed Meeting" and "Fed Minutes" were the most popular around late 2007, mid-2013, and late 2015; corresponding to the beginning of a monetary easing, the months around the "Taper Tantrum," and the departure from the zero lower bound.

[23] These are available for download from the Finance and Economic Discussion Series working-paper version of these papers: Gürkaynak, Sack, and Wright (2006) and Gürkaynak, Sack and Wright (2008), respectively.

[24] I take this from the replication materials of Nakamura and Steinsson.

real interest rates move by a similar amount several years into the term structure. The fact that *real* interest rates move is taken as evidence of monetary non-neutrality.

I slightly augment Equation (1) in order to answer the question of whether monetary policy is more or less effective when the Fed is more-accurately representing the content of its meetings:

$$\Delta i_t = \alpha + \beta\varepsilon_t + \gamma(\varepsilon_t \times \bar{\tau}_{t-1}) + \phi\bar{\tau}_{t-1} + \text{error}_t, \tag{2}$$

where $\bar{\tau}_t$ is the 12-meeting moving average of minute-transcript similarity, which has been standardized for interpretability.[25] Because the minutes of meeting $t-1$ are released between $t-1$ and t, the value of $\bar{\tau}$ at time $t-1$ reflects the prevailing level of procedural transparency at time t. The coefficient of interest is γ. With $\gamma > 0$, monetary policy has larger effects on interest rates when transparency is above its mean—a standardization of $\bar{\tau}_t$ means that γ can be interpreted as the interaction effect when procedural transparency is one standard deviation above its mean.

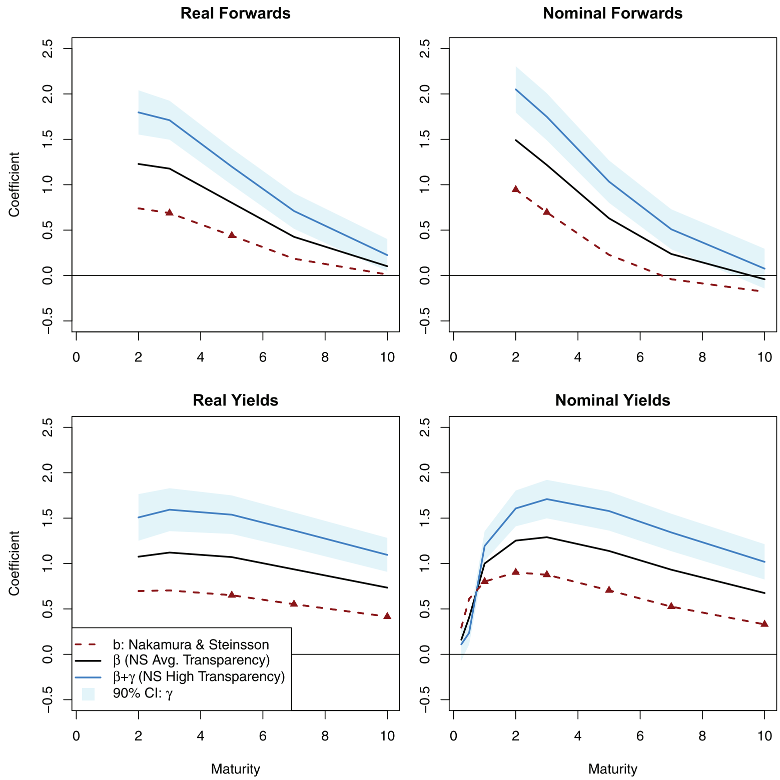
4.3 Results and Robustness

Figure 4 shows the estimated coefficients.[26] The estimated values of γ are positive for nearly every interest rate under consideration, with several of these being statistically significantly different than zero. Averaging the estimates of $\beta + \gamma$ for real yields shows that the effect of monetary policy shocks on real yields is 43 percent higher when transparency is one standard deviation above its mean than

[25]The reason for using a moving average has been discussed above. Twelve meetings is one year's worth of meetings at the beginning of the sample—this switches to eight meetings in 1981. The results are robust to using several other lag lengths—see Appendix D. Section 3 pointed to noise and observability as reasons motivating the use of a moving average—the latter is important here. If this procedural transparency is not perceptible by the public, then it is difficult to imagine reasons for which it might have effects. Given that transcripts are not released for several years, it seems unlikely that the public could realize the level of procedural transparency in real time. However, given that the Committee members give speeches and other public commentary, the public should have a sense of what is on the mind of Committee members and be able to compare this with the concerns enumerated in the minutes.

[26]Appendix E contains a table of the estimates.

Figure 4. Monetary Policy Effectiveness: Regression Results



Note: These graphs show the results of estimating Equations (1) and (2). The four panels correspond to whether i_t is a real or nominal interest rate yield or forward. The x-axis in each plot refers to the relevant maturity for each rate (3- and 6-month, and 1-, 2-, 3-, 5-, 7-, and 10-year rates). The blue lines show $\beta + \gamma$ surrounded by the 90 percent confidence interval of γ computed using heteroskedasticity-robust standard errors. The black lines show β, and the red dashed lines show b; b will be different than the corresponding values in Nakamura and Steinsson (2018), since my sample runs from January 2004 through March 2014 (in contrast to their sample of 2000–March 2014) though I follow Nakamura and Steinsson in dropping the July 2008–July 2009 period (which has little impact on the results). This forms 74 observations. The shocks are scaled so that the effect of the shock on one-year nominal yields is unity when transparency is at its average level. Transparency is standardized to have unit variance and zero mean over the sample, so that γ corresponds to the interaction effect when transparency is one standard deviation above its mean. The red triangles are drawn whenever the estimated value of b is statistically significantly below β at at least the 10 percent level, again using robust standard errors.

when it is at its average level. The estimates of β and b are consistent with the estimates and conclusion of Nakamura and Steinsson (2018)—that nominal and real rates move together far out into the term structure—even using a slightly different sample period.

The discussion in Section 3 leaned against an interpretation of minute-transcript similarity as a proxy for something unrelated to procedural transparency. That said, one might be worried that the correlation of transparency with the entropy of the transcripts (which was shown to be correlated with economic conditions) would lead the estimates of Equation (2) to simply replicate the conclusions of Tenreyro and Thwaites (2016)—namely, that monetary policy is less effective during recessions. One might also worry that transparency—having a slightly positive trend over time—is serving as a proxy for a Fed that has increasingly relied on longer-term forward guidance. Another concern is that this measure could be proxying for uncertainty regarding monetary policy. In order to alleviate these worries, Figure 5 shows the estimates of γ when different controls—and their interactions with the monetary shock—have been included in the estimating equation:

$$\begin{aligned}\Delta i_t = \alpha + \beta\varepsilon_t + \gamma(\varepsilon_t \times \bar{\tau}_{t-1}) + \phi\bar{\tau}_{t-1} \\ + \omega(\varepsilon_t \times x_t) + \pi x_t + \text{error}_t,\end{aligned} \tag{3}$$

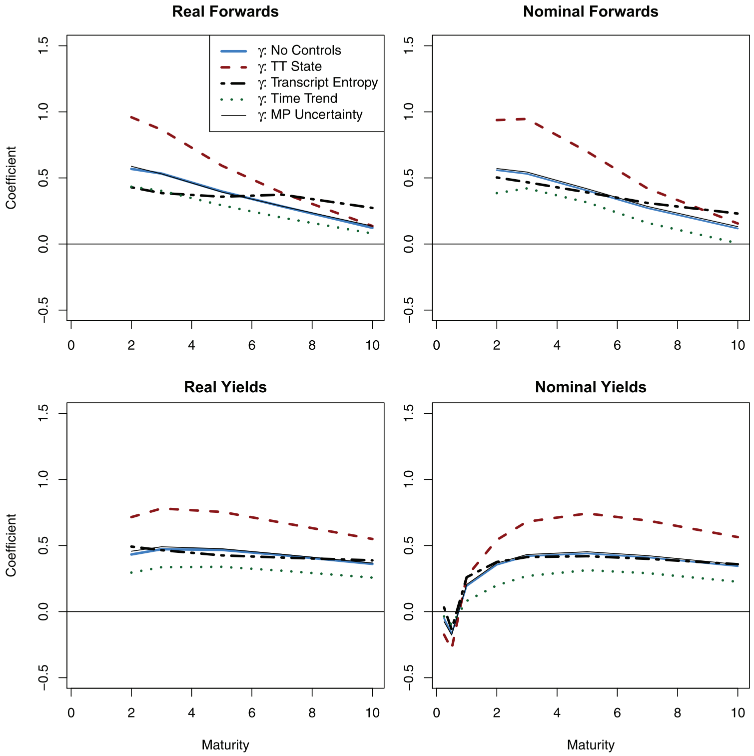
where x_t is the entropy of the transcripts, the state variable of Tenreyro and Thwaites (for the first concern), a time trend (for the second concern), or the monetary policy uncertainty index of Husted, Rogers, and Sun (2020). The results are consistent with the earlier findings—γ is positive for nearly every interest rate under consideration.

4.4 Omitted-Variables Bias

Finally, Figure 4 also highlights a slight downward bias when Equation (1) is estimated without controlling for the role that transparency plays. It is typically the case that $b < \beta$, and in a few cases this difference is statistically significant—red triangles are shown whenever $b < \beta$ is statistically significant at at least the 10 percent level.[27] Why is this the case? Mechanically, as in any

[27] This is computed using a seemingly unrelated regressions model.

Figure 5. Monetary Policy Effectiveness: Robustness of Regression Results



Note: The graphs show the coefficients γ estimated based on Equation (3). See the note to Figure 4 for details about the sample. The lines labeled "TT State" include the "state of the economy" variable of Tenreyro and Thwaites (2016) as x_t, where t refers to the quarter in which the FOMC meeting took place. The lines labeled "Transcript Entropy" have $x_t = H(\theta_{t-1})$—the lagged entropy of the transcripts. The "Time Trend" label refers to the case in which $x_t = t$. Finally, "MP Uncertainty" is the level of uncertainty about monetary policy—as measured by Husted, Rogers, and Sun (2020)—as it stood at the end of the previous FOMC meeting. Again, the monetary policy shocks are normalized so that $\beta = 1$ for the one-year nominal yields regression over the sample period.

omitted-variables bias problem, one has to consider the following relationship:

$$\varepsilon_t \times \bar{\tau}_{t-1} = \psi_0 + \psi\ \varepsilon_t + \text{error}_t.$$

The estimate of ψ will have the sign of $\text{cov}(\varepsilon_t \times \bar{\tau}_{t-1}, \varepsilon_t) \approx \mathbb{E}[\varepsilon_t^2 \bar{\tau}_{t-1}]$, which was shown to be slightly negative in Table 2—that is, monetary policy shocks tend to be larger when transparency is lower. With this relationship, however, the estimate of b will not be b but instead $\hat{b} = \beta + \gamma \cdot \psi$. With $\psi < 0$ and $\gamma > 0$, this implies that b underestimates the true effect of monetary policy shocks, β. The intuition for this result is as follows. Consider a large positive shock. The largest monetary shocks tend to occur when transparency is low ($\psi < 0$). However, low transparency also means that the effect of these large shocks on interest rates will be lower (the interaction effect, $\gamma\varepsilon_t\bar{\tau}_{t-1}$, is small or negative on average with small or negative $\bar{\tau}_{t-1}$). Theoretical work may help to clarify this chain of events, though it does suggest that transparency may be a double-edged sword for the effectiveness of monetary policy, if moving interest rates is the definition of effectiveness. Larger monetary shocks are emitted when the Fed is being less transparent about its discussions. But, possibly because the public cannot make as much sense of these short-run surprises, the shocks are less-easily transmitted to longer term interest rates. The next section provides empirical evidence in support of this interpretation.

5. Interpretation: Transparency about What?

The main transparency measure used in the estimation of Equation (2) is an aggregate measure of transparency, though the richness of the underlying data allows for transparency to be measured along different dimensions. In Section 5.1, I revisit my estimates using two new measures of transparency: transparency about discussion of the economic outlook, and transparency about monetary policy strategies. I find that my empirical estimates are driven by the latter. To provide additional color to these results I estimate, in Section 5.2, transparency regarding topics that can be used to “predict” inflation and FOMC dissents, described at the end of Section 2.3. Consistent with the results in Section 5.1, I find that transparency about monetary policy discussions—presumably contentions discussions if they predict FOMC dissents—allow interest rate shocks to pass through more fully to longer-term rates. That is not robustly the case for transparency about topics that primarily reveal information about the state of the economy.

5.1 Economic Outlook and Policy Strategy Transparency

The structure of the transcripts suggests a natural first step towards understanding what minute-transcript similarity is picking up, and which aspect of it plays an important role in determining the effectiveness of monetary policy. As documented by Hansen, McMahon, and Prat (2017), "FOMC meetings have two major parts related to the monetary policy decision: the economic situation discussion ... followed by the monetary policy strategy discussion." The authors treat these sections as separate in their analysis, and in this section I follow their lead—in so doing, I refer to the first section as ECSIT, and the latter as MPS. I use the breakdowns of the transcript and minutes in order to create an ECSIT transparency index and an MPS transparency index.[28] The ECSIT index reflects the extent to which the Committee provides details about its reading of the state of the economy to the public through the minutes. The MPS index, on the other hand, captures discussions about what this reading implies for monetary policy. The MPS therefore includes discussions of the models and targets preferred by policymakers, and their policy preferences more generally.

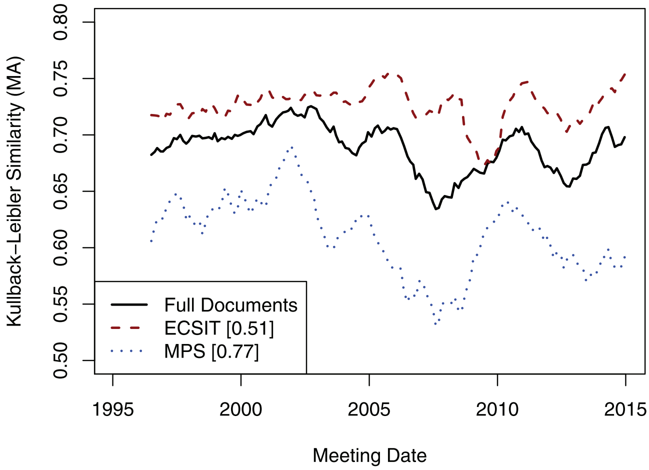
In order to create these indices, I estimate the document distributions for the ECSIT and MPS sections of the minutes and transcripts of each meeting.[29] This gives, for every FOMC meeting, a measurement of the transparency of the ECSIT portion of the meeting, and of the MPS portion, displayed in Figure 6.

Figures 7 and 8 repeat the analysis regarding the role that transparency plays in determining the effectiveness of monetary policy. Specifically, they present the results of estimating Equations (1) and (2), where $\bar{\tau}_{t-1}$ is replaced with the 12-meeting moving average of MPS and ECSIT similarity, respectively. The results for MPS similarity are quite similar to those that included minute-transcript

[28] Stephen Hansen graciously supplied the breakdown of the transcripts through 2011, and I updated this breakdown through 2014. I also performed the corresponding split for the minutes manually since 1995. To have a sense, the ECSIT portion of the minutes in 2014 included the sections titled Staff Review of the Economic Situation, Staff Review of the Financial Situation, Staff Economic Outlook, and Participants' Views on Current Conditions and the Economic Outlook, and the MPS section was the section titled Committee Policy Action. More details can be provided on request.

[29] Using the same estimated LDA topics as above.

Figure 6. Transparency by Part of Transcript



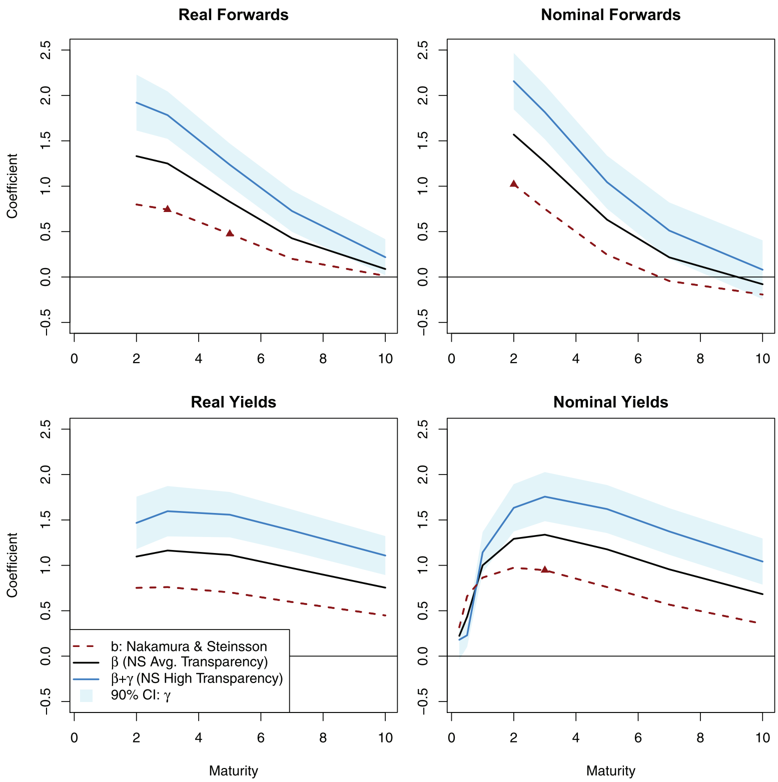
Note: This graph shows the 12-meeting moving average of transparency using the full minutes and transcripts (solid black line); the ECSIT portions of the minutes and transcripts (red dashed line); and the MPS portion of the minutes and transcripts (blue dotted lines). The bracketed numbers represent the correlation coefficient of the (moving average of the) ECSIT and MPS transparency measures with the full-document transparency measure. For reference, the ECSIT and MPS discussions take up on average 73 and 12 percent of the minutes, and 51 and 25 percent of the transcripts. These ratios have remained essentially constant since 1995.

similarity (Figure 7). This is not true for ECSIT similarity, despite the fact that ECSIT takes up a much larger portion of the FOMC discussion. Thus, the fact that monetary policy is more effective when transparency is elevated owes more to transparency regarding policymakers' views about the appropriate monetary policy—conditional on their reading of the state of the economy—than to transparency about the readings themselves.

5.2 Topic-Specific Transparency

An alternative approach for measuring transparency along different dimensions is to focus on the transparency of specific topics estimated by the LDA language model. In this section I study the role

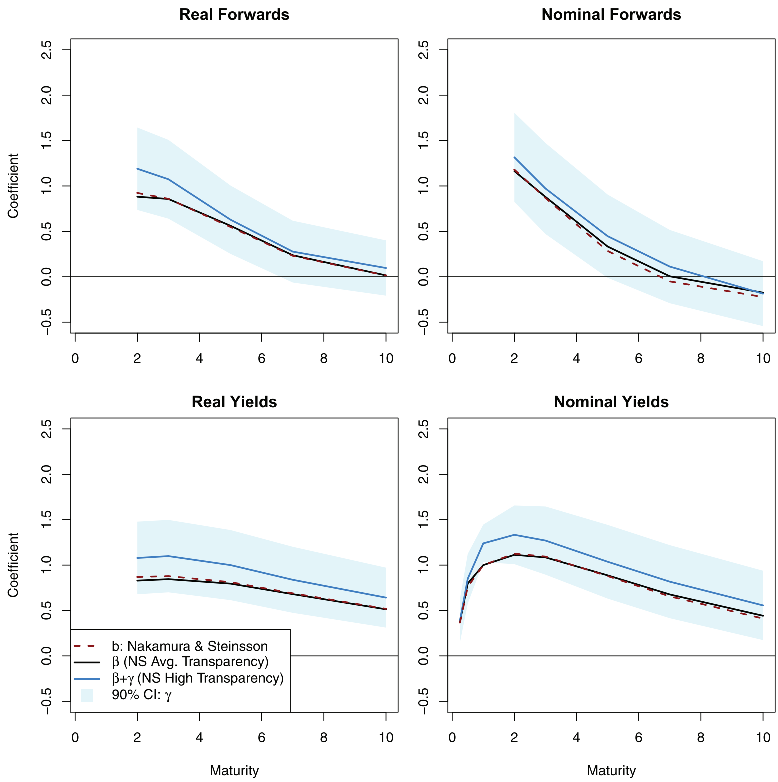
Figure 7. Monetary Policy Effectiveness: Regression Results Using MPS Similarity



Note: These graphs present the results of estimating Equations (1) and (2), where $\bar{\tau}_{t-1}$ is replaced with the 12-meeting moving average of MPS similarity. Everything else is the same as in Figure 4, so its note can be referenced for further details.

of transparency about topics that are similar to those presented in the previous section: transparency about monetary policy, and transparency about economic fundamentals. To assess the first, I measure the transparency about topics whose presence in the minutes or transcripts predict the number of dissents at each FOMC meeting in a Lasso regression. Of the 50 topics estimated by LDA, I denote the topics selected to predict dissents by $\mathcal{K}^D \subseteq \{1, \ldots, 50\}$. For transparency regarding economic fundamentals, I measure the

Figure 8. Monetary Policy Effectiveness: Regression Results Using ECSIT Similarity



Note: These graphs present the results of estimating Equations (1) and (2), where $\bar{\tau}_{t-1}$ is replaced with the 12-meeting moving average of ECSIT similarity. Everything else is the same as in Figure 4, so its note can be referenced for further details.

transparency about topics whose presence in the minutes or transcripts predict the level of CPI inflation in the month corresponding to each FOMC meeting. I denote these topics by $\mathcal{K}^{\pi}$. More details regarding the selection of these topics can be found at the end of Section 2.3.

I define the transparency for dissent-related topics τ_t^D and inflation-related topics τ_t^{π} as the cosine similarity between the (truncated) distributions of the minutes and transcripts over the selected topics. Recalling from Section 2.4 that the topic distributions of the minutes and transcripts are given by μ_t and θ_t, the transparency measures are given by

$$\tau_t^D = \frac{\sum_{k \in \mathcal{K}^D} \theta_t^k \mu_t^k}{\sqrt{\left(\sum_{k \in \mathcal{K}^D} (\theta_t^k)^2\right)\left(\sum_{k \in \mathcal{K}^D} (\theta_t^k)^2\right)}}$$

$$\tau_t^\pi = \frac{\sum_{k \in \mathcal{K}^\pi} \theta_t^k \mu_t^k}{\sqrt{\left(\sum_{k \in \mathcal{K}^\pi} (\theta_t^k)^2\right)\left(\sum_{k \in \mathcal{K}^\pi} (\theta_t^k)^2\right)}}.$$

Cosine similarities are a commonly used measure of document similarity in the NLP literature, and are simply the uncentered correlation coefficient between the topics used in each document.[30]

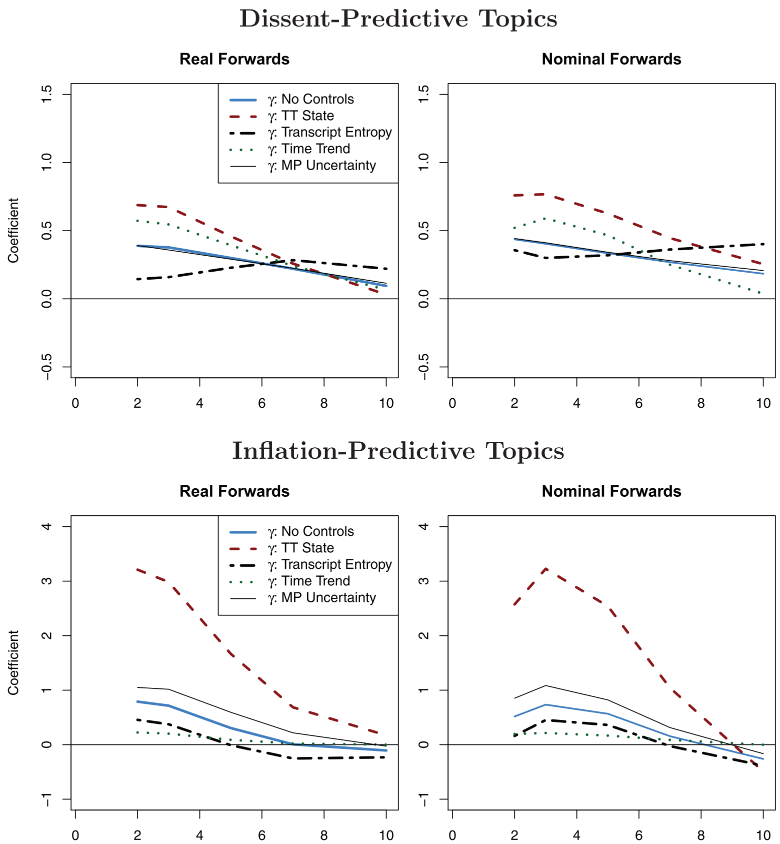
Figure 9 shows the estimates of γ for the two transparency measures for specifications that include the various controls considered in the construction of Figure 5. The estimates of γ from the specification that uses τ^π are not robustly positive, in contrast to estimates of γ estimated using dissent-predictive topics. This further suggests that it is topics relevant to the setting of monetary policy—conditional on an economic outlook—that causes changes in short-term interest rates to affect longer term nominal and real interest rates.

6. Conclusion

A central bank has a plethora of channels through which it can be transparent. Whether it is the publication of inflation reports, timely summaries of policy decisions, or post-meeting press conferences, the objective typically is to explain to the public the rationale behind policy decisions. Nowhere can these rationales be better captured than in the actual meetings in which these decisions are considered and made. The measure of transparency I propose in this paper—the similarity between the minutes and transcripts of each FOMC

[30] I use cosine similarity instead of the Kullback-Leibler divergence because the distributions of the minutes and transcripts over the selected topics do not sum to unity.

Figure 9. Monetary Policy Effectiveness: Regression Results Using Topic-Specific Similarities



Note: These graphs show the coefficients γ estimated based on Equation (3), where $\bar{\tau}_{t-1}$ is replaced with the 12-meeting moving average of the similarities of the topics in the minutes and transcripts that are predictive of the number of FOMC dissents and inflation. Everything else is the same as in Figure 5, so its note can be referenced for further details. Notice the wider axis for the inflation panel.

meeting—captures the extent to which the content of these meetings is described to the public. While this minute-transcript similarity has fluctuated over time, generally the Fed has become more transparent about its reasoning over the last 40 years. My measure is only weakly correlated with economic conditions and the policymaking environment more generally, further supporting the case that minute-transcript similarity captures something more than simply fluctuations in the discussions of FOMC meetings. Anecdotal evidence shows that this measure does indeed capture meaningful discrepancies between the two documents, and there is little evidence to suggest that these discrepancies are purposeful—they are more likely to do with the fact that the writers of the minutes face a difficult task.

Evidence suggests that monetary policy shocks have larger effects on interest rates when minute-transcript similarity is high. Additionally, when the role of transparency is neglected, these shocks seem to have smaller effects because the largest of these shocks tend to be delivered at times when transparency is low, which is also when the shocks have smaller effects on interest rates. These results suggest that high transparency allows the public to better understand what monetary policy communications and short-term interest rate movements imply for the path of future policy, captured by longer-term interest rates.

Appendix A. Transcripts, Minutes, ROPAs, MOAs, and MODs: The History of FOMC Communications

Understanding the logistical aspects of procedural transparency—which documents are released and when—is a necessary step in the assessment of procedural transparency in the "quality" sense. Since the inception of the modern-day FOMC, it has always communicated in some way the content of its meetings.

Table A.1 lists the various FOMC publications since 1935, along with their release lags.[31] While the nomenclature of the various documents has undergone several changes over the past 80 years, there have, in general, been two types of documents: detailed accounts of FOMC meetings and summaries. In general, the latter were more readily available to the public. Most of the changes in FOMC communications have occurred alongside calls for transparency, and they have formally come from congressional pressure, legislation, and litigation. The exogeneity of these external pressures permits the study of how the FOMC's procedural transparency responds to these calls.

[31] Statements, press conferences, and other releases are omitted, since the focus here is on documents whose primary purpose is to convey the meeting discussion.

Table A.1. FOMC Publications: Banking Act of 1935 to Present (2015)

Date	Meeting Summaries	Detailed Accounts
1935–67	Record of Policy Actions *(Annually)*	Minutes *(Confidential)*
1967–75	Record of Policy Actions *(90 Days)*	Memorandum of Discussion *(Five Years)*
1975–76	Record of Policy Actions *(45 Days)*	Memorandum of Discussion *(Five Years)*
1976–93	Record of Policy Actions and Minutes of Actions *(One Meeting)*	Transcripts *(Confidential)*
1993–2005	Minutes *(One Meeting)*	Transcripts *(Five Years)*
2005–	Minutes *(Three Weeks)*	Transcripts *(Five Years)*

Note: Release lags are in italics. Adapted from Danker and Luecke (2005).

This section provides an overview of the logistical aspects of these responses.

The first significant step toward greater procedural transparency—in the sense of timeliness—came in response to the 1967 Freedom of Information Act (FOIA). Beginning with the April 1967 meeting, the Record of Policy Actions—a summary of the Committee's policy actions and rationales—would be published after a 90-day lag (Danker and Luecke 2005). And, for the first time, a transcript-like document—the Memorandum of Discussion (MOD)—was to be released with a five-year lag. This set a precedent for publishing long accounts of FOMC meetings, but because the MOD was a heavily edited account, the 1967–76 period is not included in the measurement of transparency reported below.

The 1976 MODs were the last published by the FOMC; after five years of fighting a claimed FOIA violation, the Committee decided in 1981 to discontinue the MOD,[32] largely at the request of Chairman Burns (Lindsey 2003). At this point, the Committee decided to release an expanded ROPA shortly after each subsequent meeting—effectively, a 30-day lag. At the time, the reason cited for the discontinuation of the MOD was that "the benefits derived

[32]See Goodfriend (1986) for a thorough account.

from them did not justify their relatively high costs, particularly in light of the changes made in the [ROPA]" (FOMC Records of Policy Actions 1976–93, May 18, 1976). However, the more accurate reason seems to be "'fear that Congress would request access' [to the MOD] promptly" (Lindsey 2003, p. 8) and, as an FOMC subcommittee indicated, "concern about the ability to conduct monetary policy, if the court required prompt release of the memoranda of discussion" (Meltzer 2010, p. 976). The discontinuation of the MOD started a nearly 20-year period in which the FOMC published no detailed account of its meetings. Most FOMC members were aware that meetings were recorded, but they also believed that these tapes, used only for the production of minutes by Board staff, were recorded over after each meeting.

Contrary to what most members believed, congressional inquiries (primarily headed by Congressman Henry González) and internal Fed investigations revealed that, in fact, these tapes had been maintained since 1976. In November 1993, the Committee agreed to publish all of the transcripts since 1976; by 1995 the decision was made to reinstate the publication of meeting transcripts after a five-year lag. In addition, the ROPA and MOA were now combined to form the "minutes." In 2005, these minutes began to be released with a three-week lag.

All ROPAs, MOAs, minutes, and transcripts were downloaded from http://federalreserve.gov, either in PDF format or plain text. Documents in PDF format were converted to plain text using optical character recognition (OCR) software.

Appendix B. Definition and Relevance of Procedural Transparency

When used in the context of monetary policy, the word "transparency" can carry different connotations. To understand how the term is used here, Table B.1 presents the five forms of transparency relevant to central banks, as defined by Geraats (2002).[33] Procedural

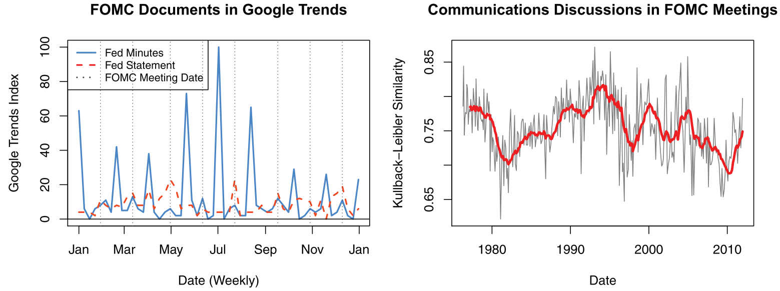
[33]Geraats has written much about central bank transparency. See Geraats (2001), where these terms were first defined, or Geraats (2007) for other examples.

Table B.1. The Types of Transparency Relevant to Central Banks (Geraats 2002, p. F540)

	Type of Transparency and Description
1.	Political transparency refers to openness about policy objectives and institutional arrangements that clarify the motives of monetary policymakers. This could include explicit inflation targets, central bank independence, and contracts.
2.	Economic transparency focuses on the economic information that is used for monetary policy, including economic data, policy models, and central bank forecasts.
3.	Procedural transparency describes the way monetary policy decisions are made. This includes the monetary policy strategy and an account of policy deliberations, typically through minutes and voting records.
4.	Policy transparency means a prompt announcement and explanation of policy decisions, and an indication of likely future policy actions in the form of a policy inclination.
5.	Operation transparency concerns the implementation of monetary policy actions, including a discussion of control errors for the operating instrument and macroeconomic transmission disturbances.

transparency—my focus in this paper—encompasses the procedure by which the accounts of FOMC decisions are released to the public via documents. What makes procedural transparency important is that increased procedural transparency presumably leads to increases in the other four types of transparency, and it is the mechanism through which the other four are manifested. For example, mandating that the Fed release the theoretical rule it uses to determine policy is a form of economic, policy, and political transparency. However, effective implementation of this policy hinges on effective procedural transparency because both the rule and deviations from it require detailed explanations. The same goes for the practice of establishing and explaining an inflation target—especially a "medium-term" target, as is done in current practice. Simply put, the communications of the Fed are "a chance for [the FOMC] to say what [they] are up to and why" (FOMC Transcripts 1976–2008, Alan Blinder, Jan. 1995).

Figure B.1. Public Interest in Fed Communications



Note: The left panel presents results from the Google Trend queries "fed statement" and "fed minutes." The y-axis represents the frequency with which a given phrase is searched on Google, and is normalized so that the highest frequency is 100. Thus, this graph does not say how these search terms rank among all other terms, but it does give information about when the terms are searched. Dashed lines are included at FOMC meeting dates. The right panel shows the Kullback-Leibler similarity of the transcripts at each point in time with a query that contains the names of FOMC minutes and transcripts over time, and the word *communication*. The full list is "transcripts minutes record policy actions memorandum discussion communications."

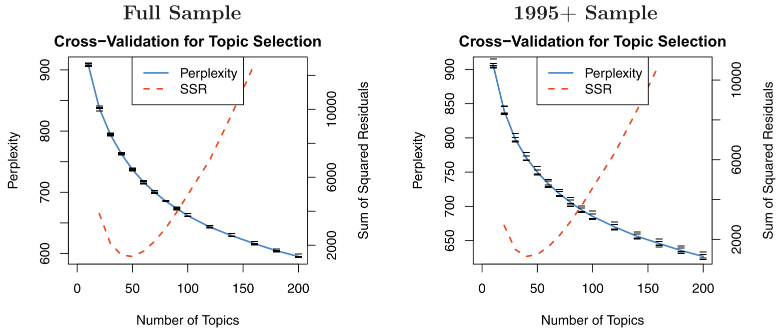
Fed communications also receive a considerable amount of attention from the public at large, and the FOMC itself. Figure B.1 supports this claim. The left panel is a graph from Google Trends—a service from Google that plots the "interest over time" of any search term—that shows how popular "fed statement" and "fed minutes" were over the year 2017. As expected, peaks in interest in the terms are in one-to-one correspondence with FOMC meetings; the three-week lagged release of minutes is also clearly noticeable. So, at the very least, there appears to be public interest in the content of FOMC documents.

The right panel of Figure B.1 provides evidence that the FOMC discusses issues of procedural transparency in its meetings. Using the text-analysis techniques described in the paper, it portrays the extent to which topics of procedural transparency were discussed at each FOMC meeting. This is measured by computing the similarity between the transcripts and a list of procedural-transparency-related words (*transcripts, minutes, communication*, etc.). The similarity of this topic with the transcripts fluctuates meaningfully, coinciding with changes in communications policy—1976 marked the temporary end of transcript publication, and topics of procedural transparency persisted for a few years after that change. Since the early 1990s, changes in publication policy have been relatively frequent—in 1993, the FOMC decided to start making its transcripts public; and in 2005, the lag between the meeting and the release of the minutes was reduced. These changes are visible on the graph, indicating that a significant amount of discussion was behind each decision. Thus, given the role of procedural transparency in monetary policymaking generally, and the fact that both the public and the Fed pay close attention to the documents used here to measure it, this paper is devoted to its study.

Appendix C. Number of Topics: Cross-Validation and Robustness

A common way to select topics for LDA is to estimate the model using a fraction of documents in the corpus, then compute how "perplexed" the model is by the held-out documents that were not used for estimation. In Figure C.1, I show the results of performing this cross-validation for different values of K (from 10–100 by 10, and from 120–200 by 20). For each value of K, the sample is split into thirds—two-thirds is used as a training sample, the term distributions are estimated, then using these distributions the topic distributions for the held-out documents are computed. The perplexity of

Figure C.1. Cross-Validation Results

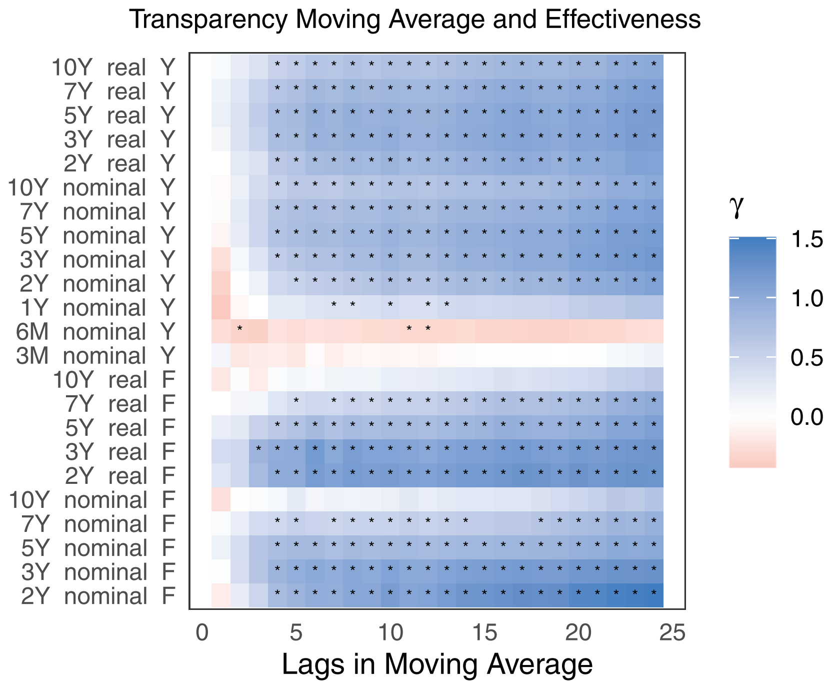


these held-out documents is then computed. This is done five times for each value of K, shown by the black lines in the figure. The blue line shows the average of the black dots for each value of K. The red dashed line shows the sum of squared of residuals from fitting two lines to the blue curve, where the two lines are split at a given value of K. This is called finding the "knee point" of the blue line, or the point at which the perplexity drops off most sharply. The knee of perplexity here is at $K = 50$, so 50 topics are used for the analysis in the paper.

Appendix D. Robustness for Moving Averages

Figure D.1 shows the results of γ from Equation (2), where the moving average of $\bar{\tau}$ uses lag lengths from 0 to 25. Stars are drawn whenever the level of statistical significance is at least 10 percent. All standardizations and scalings are performed as described in Table A.1. Notably, the coefficients are heavily skewed towards positive. Also, given how unlikely it is that the minute-transcript similarity for the most-recent meeting is observed (given that the transcript is not released for five years), the first few moving averages, while negative, should be given little weight.

Figure D.1. Estimates of γ for Different Moving Averages



Appendix E. Tabular Representation of Figures 4 and 5

Table E.1. Tabular Representation of Figures 4 and 5

	Baseline		TT State		Trans. Entropy		Time Trend		MPU	
	β	γ	β	γ	β	γ	β	γ	β	γ
Real Forwards										
2-Year	1.23	0.57	2.39	0.96	1.19	0.43	−2.00	0.43	1.28	0.59
	(0.27)	(0.15)	(1.36)	(0.27)	(0.26)	(0.18)	(1.32)	(0.14)	(0.46)	(0.14)
3-Year	1.18	0.53	2.00	0.87	1.14	0.39	−1.72	0.40	1.48	0.53
	(0.24)	(0.13)	(1.16)	(0.24)	(0.24)	(0.18)	(1.41)	(0.14)	(0.41)	(0.13)
5-Year	0.80	0.40	0.75	0.59	0.79	0.36	−0.81	0.29	1.33	0.39
	(0.22)	(0.12)	(0.88)	(0.20)	(0.21)	(0.20)	(1.26)	(0.12)	(0.34)	(0.12)
7-Year	0.43	0.28	−0.22	0.39	0.45	0.37	−0.03	0.20	0.92	0.29
	(0.21)	(0.12)	(0.80)	(0.18)	(0.20)	(0.19)	(0.95)	(0.10)	(0.32)	(0.12)
Nominal Forwards										
2-Year	1.49	0.56	2.34	0.94	1.45	0.50	−0.56	0.38	1.98	0.57
	(0.24)	(0.16)	(1.43)	(0.24)	(0.24)	(0.27)	(0.93)	(0.12)	(0.46)	(0.14)
3-Year	1.22	0.53	2.35	0.95	1.19	0.47	−1.60	0.42	1.63	0.54
	(0.27)	(0.16)	(1.41)	(0.26)	(0.26)	(0.25)	(1.39)	(0.15)	(0.48)	(0.15)
5-Year	0.63	0.40	1.13	0.70	0.63	0.39	−1.20	0.31	0.89	0.42
	(0.27)	(0.14)	(1.04)	(0.25)	(0.25)	(0.22)	(1.59)	(0.15)	(0.38)	(0.15)
7-Year	0.24	0.27	−0.01	0.42	0.25	0.31	0.24	0.16	0.48	0.28
	(0.27)	(0.13)	(0.89)	(0.24)	(0.23)	(0.23)	(1.34)	(0.13)	(0.33)	(0.15)

(continued)

Table E.1. (Continued)

	Baseline		TT State		Trans. Entropy		Time Trend		MPU	
	β	γ	β	γ	β	γ	β	γ	β	γ
Real Yields										
2-Year	1.08	0.43	1.58	0.72	1.08	0.49	−0.57	0.30	1.16	0.46
	(0.27)	(0.16)	(1.16)	(0.25)	(0.25)	(0.20)	(0.89)	(0.11)	(0.51)	(0.13)
3-Year	1.12	0.47	1.78	0.78	1.12	0.46	−1.00	0.34	1.24	0.49
	(0.25)	(0.14)	(1.16)	(0.24)	(0.24)	(0.19)	(0.98)	(0.11)	(0.47)	(0.13)
5-Year	1.07	0.47	1.61	0.75	1.06	0.43	−1.10	0.34	1.32	0.48
	(0.22)	(0.13)	(1.05)	(0.22)	(0.21)	(0.17)	(1.09)	(0.12)	(0.41)	(0.12)
7-Year	0.94	0.43	1.21	0.67	0.93	0.41	−0.88	0.31	1.26	0.43
	(0.21)	(0.12)	(0.92)	(0.20)	(0.20)	(0.17)	(1.06)	(0.11)	(0.37)	(0.11)
Nominal Yields										
3-Month	0.16	−0.05	−0.72	−0.17	0.18	0.03	0.52	−0.04	0.53	−0.08
	(0.11)	(0.11)	(0.72)	(0.14)	(0.11)	(0.12)	(0.21)	(0.06)	(0.25)	(0.08)
6-Month	0.40	−0.17	0.17	−0.28	0.40	−0.13	0.61	−0.10	0.52	−0.18
	(0.10)	(0.08)	(0.26)	(0.11)	(0.10)	(0.10)	(0.24)	(0.05)	(0.18)	(0.07)
1-Year	1.00	0.19	1.00	0.27	1.00	0.26	1.00	0.08	1.00	0.21
	(0.16)	(0.10)	(0.76)	(0.14)	(0.16)	(0.16)	(0.29)	(0.06)	(0.28)	(0.09)
2-Year	1.25	0.35	1.50	0.54	1.24	0.38	0.62	0.20	1.47	0.37
	(0.19)	(0.12)	(1.00)	(0.17)	(0.18)	(0.20)	(0.42)	(0.08)	(0.35)	(0.11)
3-Year	1.29	0.42	1.80	0.68	1.27	0.41	0.02	0.27	1.59	0.43
	(0.20)	(0.13)	(1.12)	(0.19)	(0.20)	(0.22)	(0.63)	(0.09)	(0.38)	(0.12)
5-Year	1.14	0.44	1.80	0.74	1.12	0.42	−0.63	0.31	1.45	0.45
	(0.22)	(0.13)	(1.11)	(0.21)	(0.21)	(0.22)	(0.97)	(0.11)	(0.38)	(0.12)
7-Year	0.93	0.41	1.43	0.69	0.92	0.40	−0.58	0.29	1.22	0.42
	(0.22)	(0.13)	(0.98)	(0.20)	(0.21)	(0.21)	(1.08)	(0.11)	(0.34)	(0.12)
10-Year	0.68	0.34	0.87	0.56	0.67	0.36	−0.10	0.23	0.96	0.36
	(0.21)	(0.12)	(0.84)	(0.20)	(0.20)	(0.20)	(1.05)	(0.11)	(0.30)	(0.12)

Note: The table shows the coefficients β and γ estimated based on Equation (2) (for columns labeled "Baseline") and Equation (3) (for the other columns). See the notes to Figures 4 and 5 for details about the sample, column labels, and normalizations. Robust standard errors are in parentheses.

Appendix F. Robustness to Text-Analysis Features

Figure F.1 shows the moving average of the transparency measure using LDA topics estimated over the 1995+ sample (black dashed line), which has a correlation of 0.86 with the full-sample measure. In addition, the gray solid line shows the measure with $K = 30$, which has a correlation of 0.90 with the 1995+ sample measure with $K = 50$.

Figure F.2 shows the results of estimating (2) for interest rate forwards when the LDA model is estimated on the full sample.

Figure F.1. Robustness of Transparency Measure to *K* and Sample Period

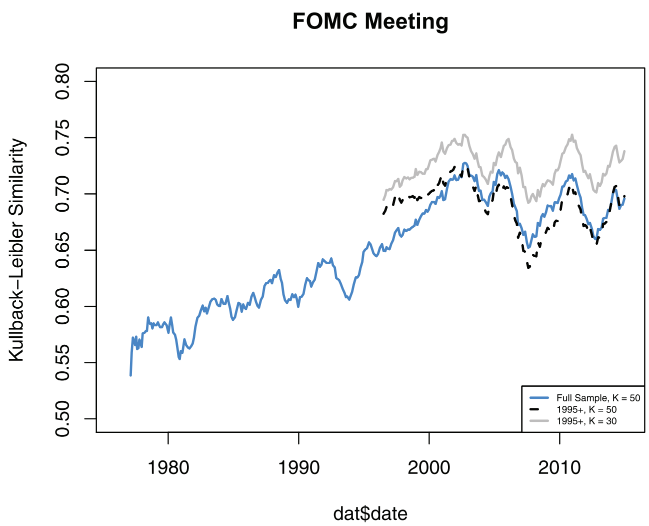
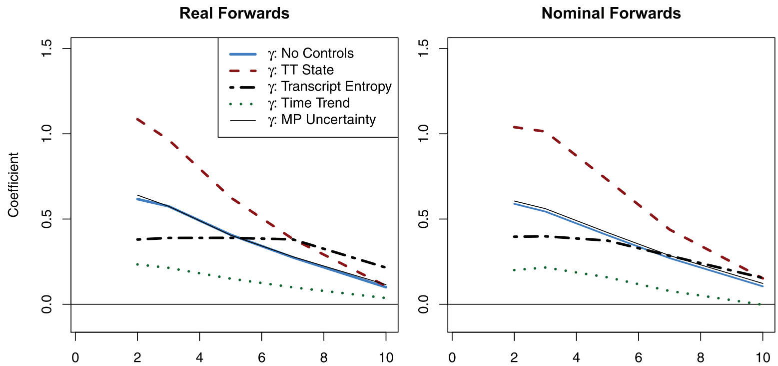
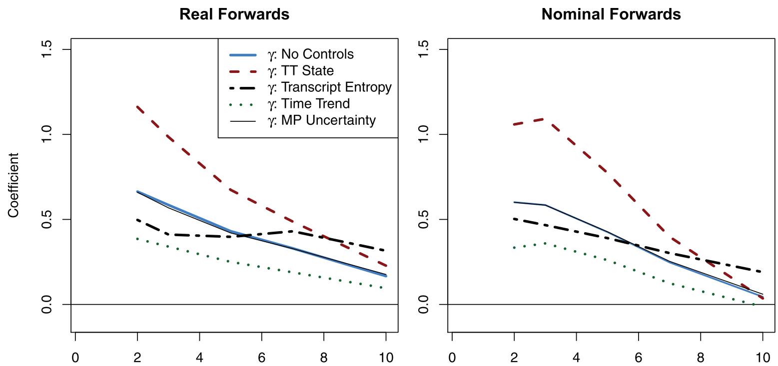


Figure F.2. Robustness of Regression Results to K and Sample Period

50-Topic, 1976–1995 Topic Model



30-Topic, 1995–2014 Topic Model



Note: The graphs show the coefficients γ estimated based on Equation (3), where $\bar{\tau}_{t-1}$ is estimated using the 1976–2014 LDA model (top panel) and the 30-topic 1995+ LDA model (bottom panel). Everything else is the same as in Figure 5, so its note can be referenced for further details.

References

Acosta, M. 2015. "FOMC Responses to Calls for Transparency." Finance and Economics Discussion Series No. 2015-60, Board of Governors of the Federal Reserve System.

Baker, S. R., N. Bloom, and S. J. Davis. 2016. "Measuring Economic Policy Uncertainty." *Quarterly Journal of Economics* 131 (4): 1593–1636.

Baum, C. F., M. E. Schaffer, and S. Stillman. 2010. "IVREG2: Stata Module for Extended Instrumental Variables/2SLS, GMM and AC/HAC, LIML and k-class Regression."

Bernanke, B. S. 2010. "Central Bank Independence, Transparency, and Accountability." Speech at the Institute for Monetary and Economic Studies International Conference, Bank of Japan, Tokyo, May 25.

———. 2013. "Communication and Monetary Policy." Speech at the National Economists Club Annual Dinner, Herbert Stein Memorial Lecture, Washington, DC, November 19.

Blei, D. M., A. Y. Ng, and M. I. Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3 (Jan): 993–1022.

Blinder, A. S., M. Ehrmann, M. Fratzscher, J. De Haan, and D.-J. Jansen. 2008. "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence." *Journal of Economic Literature* 46 (4): 910–45.

Danker, D. J., and M. M. Luecke. 2005. "Background on FOMC Meeting Minutes." *Federal Reserve Bulletin* 91 (Spring): 175–79.

Dincer, N. N., and B. Eichengreen. 2014. "Central Bank Transparency and Independence: Updates and New Measures." *International Journal of Central Banking* 10 (1): 189–259.

Egesdal, M., M. Gill, and M. Rotemberg. 2015. "How Federal Reserve Discussions Respond to Increased Transparency." https://ssrn.com/abstract=2676429.

Eijffinger, S. C. W., and P. M. Geraats. 2006. "How Transparent are Central Banks?" *European Journal of Political Economy* 22 (1): 1–21.

FOMC Records of Policy Actions. 1976–1993. Board of Governors of the Federal Reserve System.

FOMC Transcripts. 1976–2008. Board of Governors of the Federal Reserve System.

Geraats, P. M. 2001. "Why Adopt Transparency? The Publication of Central Bank Forecasts."

———. 2002. "Central Bank Transparency." *Economic Journal* 112 (483): F532–F565.

———. 2007. "The Mystique of Central Bank Speak." *International Journal of Central Banking* 3 (1): 37–80.

Goodfriend, M. 1986. "Monetary Mystique: Secrecy and Central Banking." *Journal of Monetary Economics* 17 (1): 63–92.

Griffiths, T. L., and M. Steyvers. 2004. "Finding Scientific Topics." *Proceedings of the National Academy of Sciences* 101 (suppl 1): 5228–35.

Gürkaynak, R. S., B. P. Sack, and E. T. Swanson. 2007. "Market-Based Measures of Monetary Policy Expectations." *Journal of Business and Economic Statistics* 25 (2): 201–12.

Gürkaynak, R. S., B. Sack, and J. H.Wright. 2006. "The U.S. Treasury Yield Curve: 1961 to the Present." Finance and Economics Discussion Series No. 2006-28, Board of Governors of the Federal Reserve System.

———. 2008. "The TIPS Yield Curve and Inflation Compensation." Finance and Economics Discussion Series No. 2008-05, Board of Governors of the Federal Reserve System.

———. 2010. "The TIPS Yield Curve and Inflation Compensation." *American Economic Journal: Macroeconomics* 2 (1): 70–92.

Hansen, S., M. McMahon, and A. Prat. 2017. "Transparency and Deliberation within the FOMC: A Computational Linguistics Approach." *Quarterly Journal of Economics* 133 (2): 801–70.

Husted, L., J. Rogers, and B. Sun. 2020. "Monetary Policy Uncertainty." *Journal of Monetary Economics* 115 (November): 20–36.

Klein, D., and C. D. Manning. 2003. "Accurate Unlexicalized Parsing." In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, 423–30. Sapporo, Japan: Association for Computational Linguistics.

Lindsey, D. E. 2003. "A Modern History of FOMC Communication: 1975-2002." Board of Governors of the Federal Reserve System, Washington, DC.

Loughran, T., and B. McDonald. 2011. "When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *Journal of Finance* 66 (1): 35–65.

Lustenberger, T., and E. Rossi. 2020. "Does Central Bank Transparency and Communication Affect Financial and Macroeconomic Forecasts?" *International Journal of Central Banking* 16 (2): 153–201.

Meade, E. E. 2005. "The FOMC: Preferences, Voting, and Consensus." *Review* (Federal Reserve Bank of St. Louis) 87 (2): 93–101.

Meade, E. E., N. A. Burk, and M. Josselyn. 2015. "The FOMC Meeting Minutes: An Assessment of Counting Words and the Diversity of Views." FEDS Notes, May 26, Board of Governors of the Federal Reserve System.

Meade, E. E., and D. Stasavage. 2008. "Publicity of Debate and the Incentive to Dissent: Evidence from the US Federal Reserve." *Economic Journal* 118 (528): 695–717.

Meltzer, A. H. 2010. *A History of the Federal Reserve*, Vol. 2. University of Chicago Press.

Nakamura, E., and J. Steinsson. 2018. "High Frequency Identification of Monetary Non-Neutrality." *Quarterly Journal of Economics* 133 (3): 1283–1330.

Naszodi, A., C. Csavas, S. Erhart, and D. Felcser. 2016. "Which Aspects of Central Bank Transparency Matter? A Comprehensive Analysis of the Effect of Transparency of Survey Forecasts." *International Journal of Central Banking* 12 (4): 147–92.

Newey, W. K., and K. D. West. 1994. "Automatic Lag Selection in Covariance Matrix Estimation." *Review of Economic Studies* 61 (4): 631–53.

Romer, C. D., and D. H. Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review* 94 (4): 1055–84.

Tenreyro, S., and G. Thwaites. 2016. "Pushing on a String: US Monetary Policy is Less Powerful in Recessions." *American Economic Journal: Macroeconomics* 8 (4): 43–74.

Thornton, D. L., and D. C. Wheelock. 2014. "Making Sense of Dissents: A History of FOMC Dissents." *Review* (Federal Reserve Bank of St. Louis) 96 (3): 213–27.

Woodford, M. 2001. "Monetary Policy in the Information Economy." In *Economic Policy for the Information Economy*, 297–310. Federal Research Bank of Kansas City.

———. 2005. "Central Bank Communication and Policy Effectiveness." In *The Greenspan Era: Lessons for the Future*, 399–474. Federal Reserve Bank of Kansas City.

Woolley, J. T., and J. Gardner. 2017. "The Effect of 'Sunshine' on Policy Deliberation: The Case of the Federal Open Market Committee." *Social Science Journal* 54 (1): 13–29.

Yellen, J. 2013. "Communication in Monetary Policy." Speech to the Society of American Business Editors and Writers 50th Anniversary Conference, Washington, DC, April 4.

How Do Regulators Set the Countercyclical Capital Buffer?*

Bernhard Herz and Jochen Keller
University of Bayreuth

As part of the Basel III regulatory framework, the macroprudential countercyclical capital buffer (CCyB) was introduced to mitigate the procyclicality in the financial system. National designated authorities are supposed to set the CCyB based on a "guided discretion" approach that combines rule-based and discretionary elements. We identify a CCyB puzzle, as we do not find the credit-to-GDP gap, the recommended rule-based component of the CCyB, to be crucial for buffer decisions. Instead, designated authorities appear to base their CCyB decisions in a systematic way on the discretionary elements of the framework, namely the development of house prices and non-performing loans. We also find national institutional frameworks to be relevant for CCyB policies.

JEL Codes: G01, G21, G28.

1. Introduction

In times of financial stress, the procyclical behavior of banks is likely to generate substantial negative feedback effects on the real economy. As asset prices decline, capital positions deteriorate, pressure on margins and lending standards increases, and financial institutions restrict lending to deleverage (Brunnermeier 2009). The European Systemic Risk Board (ESRB) points out that the subsequent credit shortage aggravates the economic slowdown, with negative repercussions on banks' credit portfolios (ESRB 2014). Since most banks are both creditors and debtors, network effects are likely to emerge that threaten the stability of the financial system (Brunnermeier 2009).

*We would like to thank Cyril Couaillier, Matthias Köhler, Yves Schüler, and, in particular, an anonymous referee for very helpful comments. Corresponding author: Jochen Keller, University of Bayreuth, jochen.keller@uni-bayreuth.de.

To work against such vicious circles, the countercyclical capital buffer (CCyB) was introduced as part of the global regulatory Basel III framework after a lot of preparatory work. It "is designed to help counter pro-cyclicality in the financial system. Capital should be accumulated when cyclical systemic risk is judged to be increasing, creating buffers that increase the resilience of the banking sector during periods of stress when losses materialise" (ESRB 2014). Accordingly, the CCyB should fluctuate over the financial cycle and be fully loaded at the onset of financial crises and economic downturns.

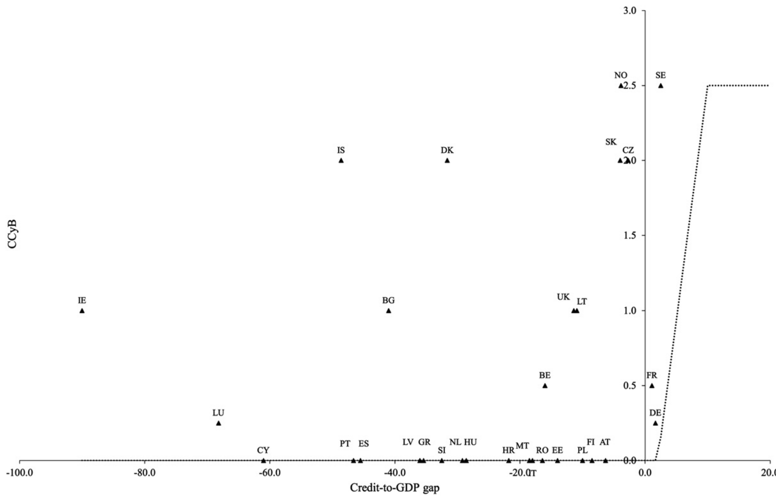
National designated authorities are supposed to implement the CCyB under a "guided discretion" approach, which combines rule-based and discretionary elements. As the rule-based component, the so-called buffer guide is based on the credit-to-GDP gap, i.e., the deviation of the credit-to-GDP ratio from its long-term trend (Basel Committee on Banking Supervision 2010; ESRB 2014). The discretionary component involves additional categories of indicators such as credit developments and private sector debt burden. These risk indicators are not specifically defined and are not subject to a specific rule so that ESRB member countries have considerable leeway in their CCyB policies.

ESRB members have used this regulatory space to a remarkable degree. On the one hand, most authorities in southern Europe (e.g., Spain, Italy, Greece, Portugal) seem to have followed ESRB recommendations and kept CCyB rates at zero, consistent with negative credit-to-GDP gaps on the national level. On the other hand, most northern European countries (e.g., Sweden, Norway, Denmark) implemented more ambitious policies and set higher CCyB rates than required by national buffer guides[1] (see Figure 1). Also, in communicating their CCyB decisions, national authorities' policies revealed a remarkable heterogeneity in how they implemented the ESRB framework on the national level.

Given that the Basel III framework has been put in place in many countries, it is time to analyze to what extent regulators actually follow these provisions. Such an analysis is particularly important given the intense discussion of the framework and the role of the

[1]Our analysis is limited to the period up to and including 2019, i.e., before the outbreak of the coronavirus pandemic. Since then, most member states have released the capital buffer.

Figure 1. Credit-to-GDP Gaps and Announced CCyB Rates



Data Source: ESRB.
Note: Latest available data as per December 2019. All values are reported in perentage points. X-axis: Credit-to-GDP gap. Y-axis: Pending CCyB. The dotted line indicates buffer guides calculated from the rule-based component.

credit-to-GDP gap as the central measure of systemic risks (see, e.g., Borio et al. 2010; Gischer, Herz, and Menkhoff 2019).

Given the wide gap between the Basel III and ESRB recommendations on the one hand and the actual CCyB policies in EU member states, on the other hand, we are interested in the key motives for national CCyB decisions. We contribute to the sparse literature on the CCyB instrument by empirically analyzing the actual drivers of CCyB decisions in European countries. In this analysis, we differentiate two dimensions of the CCyB, which are related but might be driven by different determinants. First, we address in a qualitative analysis whether or not national designated authorities make use of the countercyclical buffer. Second, we analyze the factors driving CCyB decisions over time. Both approaches provide interesting complementary information in order to better understand macroprudential policies in the EU.

In contrast to its prominent role in the ESRB (2014) recommendation, we do neither find robust empirical evidence that the credit-to-GDP gap systematically drives the buffer activation nor its variation over time, as the coefficients of the credit-to-GDP gap are not significantly different from zero. We also test the hypothesis of designated authorities following the rule-based component of the ESRB (2014) recommendation, where the CCyB is calibrated to the credit-to-GDP gap. This alternative null hypothesis is clearly rejected. Interestingly we also do not find the selected buffer guides to be crucial for CCyB decisions.

In contrast, higher house price growth and lower non-performing loans ratios make the use of the countercyclical buffer more likely. We also find evidence that developments in house prices and credit quality are relevant for CCyB adjustments over time. Thus, additional risk indicators appear to be more relevant for CCyB decisions than the credit-to-GDP gap.

Consistent with Edge and Liang (2020), we find that the institutional role of the designated authority matters. The likelihood of using the CCyB is smaller if the existing prudential regulator or the central bank takes the final decision about the buffer. In contrast, the announcement of a positive countercyclical buffer is more likely if the domestic Financial Stability Committee (FSC) is the decisionmaker.

In line with the literature, we argue that the weak relationship between the credit-to-GDP gap and actual CCyB decisions is a major challenge for the communication and the acceptance of the macroprudential instrument.

We do not claim that the credit-to-GDP gap is not considered at all by national authorities. However, it does not seem to be systematically taken into account in decisionmaking. Against the background of its highlighted importance the gap takes in official recommendations and European legislation, the results raise the question of whether the indicator is suitable for setting the buffer at all.

The remainder of this paper is organized as follows: Section 2 discusses the concept of "guided discretion" as implemented in the CCyB context and reviews the literature. Section 3 presents the data used in our empirical investigation. In Section 4, we discuss our model selection and the results of the logit and linear panel regression. Section 5 provides several robustness checks. Finally, Section 6 concludes.

2. Guided Discretion

To stabilize the financial sector, the ESRB requires designated authorities to impose a capital buffer on credit institutions and relevant investment firms (Directive 2013/36/EU 2013) based on a "guided discretion" approach that combines rule-based and discretionary elements. This CCyB rate ranges from 0 percent to 2.5 percent of risk-weighted assets (RWA), in steps of at least 0.25 percentage point. As the rule-based element, the so-called benchmark buffer rate requires a 0 percent capital buffer for credit-to-GDP gaps below 2 percentage points, a linearly increasing rate ranging from 0 percent to 2.5 percent for credit-to-GDP gaps between 2 percentage points and 10 percentage points, and a top 2.5 percent CCyB rate if the corresponding ratio is more than 10 percentage points above its long-term trend (ESRB 2014) (see Equation (1) and Figure 1).

$$\text{Benchmark buffer rate}_t(\%) = \begin{cases} 0 & if \quad Gap_t \leq 2pp \\ 0.3125 * Gap_t - 0.625 & if \quad 2pp < Gap_t < 10pp, \\ 2.5 & if \quad Gap_t \geq 10pp \end{cases} \tag{1}$$

Concerning the discretionary component, the ESRB (2014) suggests complementing the credit-to-GDP gap with several additional variables[2] to gauge the buildup of systemic risk:

(a) potential overvaluation of property prices

(b) credit developments

(c) external imbalances

(d) strength of bank balance sheets

(e) private sector debt burden

[2]Among Bank for International Settlements (BIS) member states, designated authorities in Germany take into account the largest number of core systemic risk indicators in their CCyB decisions, followed by France and the United Kingdom (BIS 2017). For an extensive discussion of the forecasting quality of the different indicators, see Detken et al. (2014) and Tölö, Laakkonen, and Kalatie (2018).

(f) potential mispricing of risk

(g) model-based risk measures that combine the credit-to-GDP gap and a selection of the above-mentioned variables.

The concept of "guided discretion" is thus specified as "a rules-based approach with the exercise of their discretionary powers when deciding on the appropriate buffer rate" (ESRB 2014). Although there is scope for national authorities, the credit-to-GDP gap is formally by far the most important indicator. As the only indicator, the gap is directly and explicitly converted into a buffer guide value (ESRB 2014). Furthermore, to improve transparency, EU legislation requires national institutions to quarterly publish the credit-to-GDP ratio, the credit-to-GDP gap, and the buffer guide (Directive 2013/36/EU 2013, Article 136). In contrast, the ESRB does not impose specific guidelines on how to account for the seven other categories of risk indicators. It is only recommended to publish variables from categories (a) to (f) if they are relevant and available (ESRB 2014).

Obviously, a necessary condition for a rule-based CCyB framework is the credit-to-GDP gap to be a good predictor of financial crises. Borio and Lowe (2002b) identify the credit-to-GDP gap as the best single indicator among a wide variety of alternative variables. Borio et al. (2010) document for a set of developed countries that pronounced above-trend increases in the credit-to-GDP ratio, i.e., positive credit-to-GDP gaps, typically precede financial crises.

When calculating credit-to-GDP ratios, two elements turned out to be of particular importance: the definition of credit and the trend extraction method to filter out the cyclical component. According to the official recommendation, national designated authorities are supposed to use a "broad measure of the stock of credit" (ESRB 2014) for computing the credit-to-GDP ratio. Drehmann (2013) uses total credit to the non-financial sector and bank credit for calculating the credit-to-GDP gap. While both aggregates are helpful in constructing early-warning tools, he finds the credit gap based on total non-financial-sector debt, which is also used in the so-called standardized credit-to-GDP gap, to better reflect the underlying risk preceding financial crises.

On a more technical level, calculating credit-to-GDP gaps involves a number of crucial assumptions on how to decompose the time series into cyclical and trend components. Borio et al. (2010) recommend a high smoothing parameter when estimating the trend of the credit-to-GDP ratio by using a one-sided (i.e., recursive) Hodrick-Prescott (HP) filter to reflect the longer duration of credit cycles compared with business cycles. In particular, they estimate the median of credit cycles to be about 15 years, and therefore, three to four times longer than standard business cycles. Under such a long duration, the corresponding smoothing parameter for quarterly data should be in the range between 125,000 and 400,000 (Borio et al. 2010). The ESRB (2014) follows this literature in recommending a one-sided HP filter with large smoothing parameter ($\lambda = 400{,}000$).

In contrast, Edge and Meisenzahl (2011) find that credit-to-GDP gaps are not a reliable basis for determining CCyB rates. In particular, volatile end-of-sample trend estimates may lead to distortions when assessing credit gaps in real time, and thus, might lead to potential ex post revisions of the gap. The critique is related to the more general observation that HP filters are plagued by spurious dynamics. Hamilton (2018) advises to refrain from using HP filters completely and to use linear projections based on the four most recent values. In contrast, Drehmann and Yetman (2018) recommend the use of HP filters when estimating credit gaps, as none of the considered alternative indicators, i.e., gaps based on linear projections and 20-quarter growth rates, systematically outperform the standard credit-to-GDP gap.

Galán (2019) regards the smoothing parameter of the standardized credit-to-GDP gap as unrealistically high since he estimates the financial cycle in most European countries to be shorter. The resulting high degree of inertia implies that the standardized gap is a biased signal for the true state of the financial cycle, with recent credit gaps remaining in deeply negative territory. There is more support for using smaller and/or more adjusted smoothing parameters (e.g., Kauko and Tölö 2019; Wezel 2019). Reigl and Uusküla (2018) investigate, in particular, the weaknesses of the standardized credit-to-GDP gap. Short time series intensify exceptionally small (i.e., negative) standardized credit gaps so that in some cases, even a pronounced credit boom would not have closed the negative gap (Reigl and Uusküla 2018).

Wolf, Mokinski, and Schüler (2020) find considerable differences between standard one-sided HP filters and their corresponding two-sided version. One-sided filters suppress higher-frequency volatility more, which is what should be extracted by the filter. They advise against the standard one-sided HP filter for extracting cyclical trends in real time and propose a lower smoothing parameter together with a multiplicative rescaling factor for the cyclical component (Wolf, Mokinski, and Schüler 2020).

As the credit-to-GDP gaps in 2019 (Figure 1 and Equation (1)) imply, buffer benchmark rates have been zero or very small in most countries. Not surprisingly, the widespread practice of designated authorities to deviate from the benchmark buffer rate has led to an intensive discussion of the ESRB recommendation.

Couaillier, Idier, and Scalone (2019) and the ESRB (2019, 2020) emphasize that some national authorities follow more ambitious CCyB policies either by applying more demanding buffer guides or explicitly accounting for additional indicators besides the credit-to-GDP gap. For instance, the United Kingdom, the Czech Republic, and Lithuania have implemented a positive "neutral rate," i.e., a positive CCyB rate even when risk is considered to be only moderate (ESRB 2019, 2020).

In the communiqués that accompany and explain CCyB decisions, national designated authorities provide further insights into their strategies and, in particular, the specific role of rule-based and discretionary elements in their CCyB policies. The Swedish Financial Supervisory Authority, e.g., declares to place "little weight on the buffer guide as an indicator to raise the buffer since the underlying trend in lending in relation to GDP deviates significantly from a level that is sustainable in the long run. Other authorities with responsibility for macroprudential tools also place little weight on the buffer guide and look at other indicators" (Finansinspektionen 2018). The BaFin (2019), as Germany's designated authority, mentions three risk categories, namely economic risk, real estate risk, and interest rate risk, by citing the recommendation of the domestic Financial Stability Committee when activating the CCyB in 2019. The BaFin (2019) further concludes that additional variables mentioned in ESRB (2014) signal the buildup of cyclical risk, e.g., developments in real estate prices, growth in housing loans, and credit growth to non-financial corporations. When activating the CCyB, the Czech National Bank (2015) indicated that the credit-to-GDP gap is not fully suitable for CCyB rate decisions in the Czech Republic and that it takes into account other indicators that better reflect the so-called converging economy. The decision to increase the buffer is primarily justified by increased credit growth. Moreover, the debt-to-income ratio, credit standards, and the property markets are also mentioned as important factors (Czech National Bank 2015).

Not so surprisingly, national decisionmakers whose capital buffer decisions are more in line with the buffer benchmarks also give more weight to the credit-to-GDP gap in explaining their buffer decisions. For instance, the Banca d'Italia (2019) vindicated its decision to leave the CCyB unchanged at 0 percent with the standardized and the nationally adjusted credit-to-GDP gap, both of which were in negative territory. In the further step, other indicators are discussed, such as the growth of bank loans, non-performing loans, and the unemployment rate. Similarly, the Banco de Portugal (2019) in its decision to leave the CCyB unchanged at 0 percent firstly addressed the standardized and the nationally adjusted credit-to-GDP gap and then discussed additional indicators, most of which sent similar signals. In doing so, the national designated authority followed the categories recommended by ESRB (2014) and explained recent developments in credit growth, credit demand and spreads, house prices, the loan-to-deposit ratio, the debt-service-ratio, and the current account balance.

In their policy evaluation Babić and Fahr (2019) discuss how positive CCyB rates in a negative credit gap environment have created major communication challenges for national macroprudential authorities. They find that the credit-to-GDP gap has only a limited impact on CCyB decisions in European countries, as national decisionmakers rely on alternative measures to identify the state of the financial cycle, e.g., a composite indicator as in Slovakia. As a result, they advocate using additional risk measures consistently. In a rare study of the role of the institutional supervisory framework for CCyB decisions, Edge and Liang (2020) find that institutionally stronger FSCs are associated with a higher likelihood of positive CCyB rates. Their analysis also indicates that the credit-to-GDP gap is not systematically relevant for CCyB decisions.

Given this evidence that the rule-based component of the regulatory framework is only of a minor, if any, relevance for CCyB decisions, the question arises of what actually drives buffer decisions in Europe. To the best of our knowledge, we are only aware of one study that empirically analyzes CCyB decisions. While Edge and Liang (2020) focus on how the institutional design of FSCs affects the initial use of the CCyB, they also control for other economic and financial indicators.[3] They find that most FSCs have relatively weak tools and seem to be motivated by symbolic delegation, i.e., signaling action to the public. The credit-to-GDP gap does not significantly affect the probability of setting positive CCyB rates (Edge and Liang 2020).

3. Data

As the ESRB provides the framework for national CCyB decisions,[4] we build on the ESRB data set and analyze CCyB policies during the time period between 2014, when the CCyB framework was implemented, to the end of 2019, the time up to the coronavirus pandemic. If there was more than one decision for a particular quarter and country, we kept the last decision. Our panel is unbalanced since designated authorities started to report CCyB decisions at different points in time. If available, the standardized credit-to-GDP data[5] are used in our analysis. In some cases, only measures calculated from narrower aggregates were reported. We include the 30 European countries from the ESRB data set (Table 4) except Norway, Iceland, and Greece, as comparable data on credit and house price developments were not available.

The ESRB (2014) mentions several complementary risk categories that might indicate the buildup of systemic risk. As additional indicators (see Table 1 for further details), we approximate the

[3]Earlier work focuses on a broader set of prudential tools as in Cerutti et al. (2016) and Cerutti, Claessens, and Laeven (2017), while experience with the CCyB in Europe was very limited. For macroprudential policies in general, Cerutti, Claessens, and Laeven (2017) and Akinci and Olmstead-Rumsey (2018) analyze the effectiveness of various macroprudential tools.

[4]In this paper, we concentrate on the announced (pending, future) CCyB, which has to be fulfilled at the end of the transitional period, which is usually one year. In between, the announced requirement may be different from the effective capital requirement. In this context, the terms "announced," "pending," and "future" are used interchangeably.

[5]We cross-checked ESRB data with data available from national macroprudential/designated authorities and corrected obvious errors.

Table 1. Variable Description

Variable	Description	Source
CCyB	Announced (pending) rate of the countercyclical capital buffer (in percent of RWA), quarterly, linearly interpolated in case of missing data.	ESRB
$CCyB^{>0}$	Indicator variable that equals 1 if $CCyB > 0$ and 0 otherwise, derived from CCyB, quarterly.	ESRB
Credit-to-GDP Gap	Credit-to-GDP gap (deviation of the Credit-to-GDP ratio from its long-term trend), quarterly, linearly interpolated in case of missing data.	ESRB, national authorities
Credit Growth (1Y)	Year-on-year growth rate (in percent) of debt securities and loans of the private non-financial sector, quarterly.	ECB
MFI Credit Growth (1Y)	Year-on-year growth rate (in percent) of MFI credit (loans and debt securities) granted to (domestic) non-financial corporations and households, quarterly.	ECB
Credit-to-GDP Ratio	Credit-to-GDP ratio, quarterly, linearly interpolated in case of missing data.	ESRB, national authorities
Buffer Guide	CCyB guide, quarterly, linearly interpolated in case of missing data.	ESRB
House Prices (5Y)	House price index growth (total, 2015 = 100) over five years, quarterly.	Eurostat
Stock Index (1Y)	Domestic stock market index, year-on-year growth rate of the quarterly mean of daily levels.	Datastream
Stock Index Volatility	Realized index volatility (in logs), calculated from daily stock market index levels, quarterly.	Datastream, own calculations
Current Account	Current account (in percent of GDP), quarterly.	Eurostat
Regulatory Capital	Regulatory capital (in percent of RWA), quarterly, linearly interpolated in case of missing data.	IMF
Non-performing Loans	Non-performing loans (in percent of total gross loans), quarterly, linearly interpolated in case of missing data.	IMF
PR Sets CCyB	Indicator variable, equals 1 if the prudential regulator sets CCyB and 0 otherwise.	ESRB, Edge and Liang (2020)
CB Sets CCyB	Indicator variable, equals 1 if the central bank sets CCyB and 0 otherwise.	ESRB, Edge and Liang (2020)
MF Sets CCyB	Indicator variable, equals 1 if the Ministry of Finance sets CCyB and 0 otherwise.	ESRB, Edge and Liang (2020)
FSC Sets CCyB	Indicator variable, equals 1 if the Financial Stability Committee sets CCyB and 0 otherwise.	ESRB, Edge and Liang (2020)

potential overvaluation of property prices (a) by the growth rate of the domestic house price index over five years. Even though changes in house prices may be fundamentally justified, real estate prices can add useful information for predicting financial crises (see, e.g., Borio and Lowe 2002a). The Basel Committee's member countries consider house price growth after credit-to-GDP measures most often for setting the CCyB (BIS 2017). Accounting for property prices in macroprudential decisions is also in line with Borio (2014), who identifies real estate prices as key drivers for the financial cycle. Moreover, house price index data for European countries is typically available on a quarterly basis with a relatively short time lag. To monitor credit developments (b), we consider the year-on-year growth rate of private non-financial-sector debt securities and loans. Even though the credit-to-GDP gap is positively correlated with credit growth, some countries exhibit substantial growth rates in debt while having negative credit-to-GDP gaps. We take quarterly current account data (in percentage of GDP) as a measure for external imbalances (c). To proxy the strength of bank balance sheets (d), we employ both regulatory capital (in percentage of RWA) and non-performing loans (in percent of total gross loans). To measure the private sector debt burden (e), the ESRB (2014) and some national supervisors propose debt-service ratios (Tente et al. 2015). Due to data limitations, we cannot take these into account. To account for potential mispricing of risk (f), we incorporate the year-on-year growth rate of the leading domestic stock market index and the corresponding realized volatility. To have comparable indicators of the domestic stock market volatilities, we calculate the volatility proxy from the quarterly sum of daily squared returns.[6] The ESRB (2014) proposes real equity price growth as a potential variable to measure the mispricing of risk. As pointed out by Tente et al. (2015), strong and sudden price increases in stock markets may indicate that risks are not correctly priced by the market. A number of studies (e.g., Detken et al. 2014; Tölö, Laakkonen, and Kalatie 2018) found

[6]In more detail, we follow Christiansen, Schmeling, and Schrimpf (2012) in defining the realized volatility as $RV_{it} = ln\sqrt{\sum_{s=1}^{Q_t} r_{its}^2}$, where Q_t denotes the number of return observations in quarter t.

that equity price developments add useful information, in particular in multivariate signaling approaches. Analogously, relatively low equity price volatility may indicate that stock investors underestimate the associated risk (Tente et al. 2015) and may lead to elevated risk-taking (Tölö, Laakkonen, and Kalatie 2018).

In addition to these macroeconomic and financial variables, we consider several institutional variables to control for differences in national regulatory governance. In investigating the decision to use the CCyB actively, we add indicator variables mirroring the role of the decisionmaker, as proposed in Edge and Liang (2020). The dummy variable "PR sets CCyB" equals one if the prudential regulator sets the CCyB and zero otherwise. Accordingly, the variables "CB sets CCyB," "MF sets CCyB," and "FSC sets CCyB" account for the central bank, the ministry of finance, and the FSC as decisionmakers. The FSC consists of multiple institutions and generally includes the central bank, the prudential regulator, and the government (Edge and Liang 2020). While the committee is the designated authority in a few cases, it has only an advisory role in most member countries. Edge and Liang (2020) show that the focus of existing institutions, e.g., financial soundness on the individual level for the prudential regulator, influences macroprudential decisions. As these institutional variables vary only between countries, but not over time in our estimation period, country fixed effects absorb their influence in the linear panel regression. Table 1 describes the time series, transformations, and raw data sources. Table 2 provides summary statistics and Table 3 coefficients of correlations for the transformed time series.

4. Estimation

There are several challenges when empirically investigating CCyB policies. First, the framework of this macroprudential tool has been implemented only recently, and many countries have not actively used the countercyclical buffer yet. Second, the dependent variable CCyB is truncated with a lower bound of CCyB $= 0\%$ and an upper bound at CCyB $= 2.5\%$. Third, the mixture of different starting points of the CCyB reporting and diverse financial structures implies an unbalanced panel in which unobserved heterogeneity is likely to be present.

Table 2. Independent Variables: Summary Statistics

Variable	Abbreviation	Mean	Std. Dev.	Min.	Max.	Observations
Credit-to-GDP Gap	Gap	–20.24	19.85	–93.00	13.50	493
Credit Growth (1Y)	CG	2.89	4.29	–8.66	44.98	493
MFI Credit Growth (1Y)	MFI CG	2.13	5.29	–29.78	13.63	454
Credit-to-GDP Ratio	Ratio	131.65	60.29	36.20	359.00	493
House Prices (5Y)	HP	20.65	19.56	–24.32	92.99	493
Stock Index (1Y)	SI	4.82	14.43	–26.81	51.89	493
Stock Index Volatility	SIV	–2.76	0.41	–3.96	–1.64	493
Current Account	CA	1.60	6.98	–45.50	29.90	493
Regulatory Capital	RC	19.84	3.48	12.27	36.08	493
Non-performing Loans	NPL	5.74	6.52	0.36	47.75	493
PR Sets CCyB	PR	0.23	0.42	0.00	1.00	493
CB Sets CCyB	CB	0.61	0.49	0.00	1.00	493
MF Sets CCyB	MF	0.08	0.27	0.00	1.00	493
FSC Sets CCyB	FSC	0.08	0.27	0.00	1.00	493

Note: Further details on data calculation and sources are provided in Table 1.

Table 3. Correlations of Explanatory Variables

Variable	Gap	CG	MFI CG	Ratio	HP	SI	SIV	CA	RC	NPL
Gap	1.000									
CG	0.434	1.000								
MFI CG	0.575	0.483	1.000							
Ratio	–0.388	–0.248	–0.459	1.000						
HP	–0.071	0.249	0.314	–0.012	1.000					
SI	–0.008	0.141	0.052	–0.078	0.118	1.000				
SIV	0.119	–0.070	–0.056	0.220	–0.153	–0.182	1.000			
CA	0.014	–0.003	0.080	0.034	0.079	0.032	–0.078	1.000		
RC	0.042	0.128	0.110	0.118	0.382	0.116	–0.189	0.025	1.000	
NPL	–0.373	–0.347	–0.543	0.351	–0.441	–0.125	–0.020	–0.204	–0.278	1.000

Note: Further details on data calculation and sources are provided in Table 1.

Table 4. Countries and Domestic CCyB Rates

Country	Code	Decision Date	CCyB (Pending Rate)
Austria	AT	2019-09-05	0.00
Belgium	BE	2019-09-16	0.50
Bulgaria	BG	2019-09-17	1.00
Croatia	HR	2019-09-30	0.00
Cyprus	CY	2019-09-10	0.00
Czech Republic	CZ	2019-08-29	2.00
Denmark	DK	2019-10-01	2.00
Estonia	EE	2019-09-30	0.00
Finland	FI	2020-09-27	0.00
France	FR	2019-07-09	0.50
Germany	DE	2019-09-30	0.25
Greece	GR	2019-09-16	0.00
Hungary	HU	2019-09-24	0.00
Iceland	IS	2019-10-01	2.00
Ireland	IE	2019-07-04	1.00
Italy	IT	2019-09-17	0.00
Latvia	LV	2019-10-29	0.00
Lithuania	LT	2019-09-27	1.00
Luxembourg	LU	2019-10-01	0.25
Malta	MT	2019-10-01	0.00
Netherlands	NL	2019-09-24	0.00
Norway	NO	2019-09-19	2.50
Poland	PL	2019-09-23	0.00
Portugal	PT	2019-10-01	0.00
Romania	RO	2019-09-11	0.00
Slovakia	SK	2019-10-21	2.00
Slovenia	SI	2019-11-05	0.00
Spain	ES	2019-09-20	0.00
Sweden	SE	2019-10-24	2.50
United Kingdom	UK	2019-10-02	1.00

Source: ESRB. Latest available data as per December 2019.

As discussed above, we differentiate between the decisions to actively use a CCyB, i.e., to announce a non-zero rate, and to set a specific level of the buffer. Obviously, the second decision is contingent on the first.

To examine the first question, i.e., the decision to activate the CCyB, we estimate a random-effects logit model as in Edge and Liang (2020),

$$Pr(CCyB_{it}^{>0}) = \frac{1}{(1 + exp[-(\alpha + \mathbf{x}_{it}'\boldsymbol{\beta} + \mathbf{z}_i'\boldsymbol{\gamma} + \delta_i)])}, \quad (2)$$

where $CCyB_{it}^{>0}$ equals one if the buffer is active with a positive announced rate for country i in quarter t and zero otherwise. $\mathbf{x}_{it}$ represents the vector of economic and financial indicators as discussed in the previous section, and $\mathbf{z}_i$ the country-specific indicator variables for the decisionmaker. Finally, δ_i denotes the unobserved effect.

As we investigate whether the capital buffer is above zero for a given country and point in time, the dependent variable varies over time and country, in contrast to Edge and Liang (2020) who only examine whether the macroprudential instrument is used or has been used for a given country. Furthermore, we estimate the model based on quarterly data instead of annual data, with missing data being replaced by linear interpolations, if necessary.

Table 5 reports the random-effects logit regression results. We do not find reliable empirical evidence for a substantial role of the credit-to-GDP gap for CCyB policies in Europe. This is obviously at odds with the prominent role of the rule-based component in the ESRB recommendation. It also reflects the weak relationship of the credit-to-GDP gap and the buffer rate, as displayed in Figure 1.

To better understand the guided discretion approach proposed by the ESRB, we examine the (non) role of the credit-to-GDP gap in greater detail. As specified by Equation (1), designated authorities are expected to activate the CCyB as soon as the gap equals 2 percentage points, making this value a pivotal point. For this 2 percentage points value of the credit-to-GDP gap, we test if an increase in the gap leads to higher predicted probabilities of positive CCyBs. Additional indicators mentioned by the ESRB recommendation may be relevant for the calibration of the CCyB, which thus may also affect the probability of its implementation. Therefore, we test against different changes in the predicted probability. We perform χ^2-tests to test the null hypothesis that the conditional marginal effect of a one-unit increase in the credit-to-GDP gap on

Table 5. Random-Effects Logistic Regression

$CCyB^{>0}$	I	II	III	IV
Credit-to-GDP Gap	0.0870	0.1061*	0.0932	0.0857
	(0.0595)	(0.0574)	(0.0736)	(0.0582)
Credit Growth (1Y)	–0.2268	–0.2211	–0.2577	–0.2219
	(0.1489)	(0.1519)	(0.1645)	(0.1514)
House Prices (5Y)	0.3360***	0.3594***	0.3868***	0.3567***
	(0.0472)	(0.0461)	(0.0558)	(0.0460)
Stock Index (1Y)	–0.0488	–0.0451	–0.0464	–0.0470
	(0.0349)	(0.0352)	(0.0377)	(0.0350)
Stock Index Volatility	0.8988	1.1637	1.1349	1.0075
	(1.2663)	(1.3089)	(1.3855)	(1.2878)
Current Account	0.0020	0.0041	0.0079	0.0002
	(0.0549)	(0.0561)	(0.0611)	(0.0542)
Regulatory Capital	0.1511	0.1032	0.1604	0.1108
	(0.2670)	(0.2710)	(0.2746)	(0.2643)
Non-performing Loans	–2.7428***	–2.8124***	–3.1334***	–2.7325***
	(0.4865)	(0.5004)	(0.6435)	(0.4674)
PR Sets CCyB		–12.6473	–8.9488**	
		(10.1847)	(3.7879)	
CB Sets CCyB		–6.8194		
		(9.2117)		
MF Sets CCyB		–7.7255		
		(10.4509)		
FSC Sets CCyB				9.5632
				(9.3317)
Observations	493	493	493	493
Log-Likelihood	–78.19	–77.51	–77.68	–77.70
χ^2 (DF)	113.26 (8)	132.43 (11)	74.75 (9)	134.60 (9)

Note: The dependent variable is the binary CCyB decision. The models were estimated using a constant, which is not reported. Significance levels: ***1 percent, **5 percent, *10 percent. Data sources are provided in Table 1.

the predicted probability $\Pr(CCyB^{>0} = 1)$ equals 0, 25, 50, 75, or 100 percentage points, respectively. Table 6 reports conditional marginal effects (based on the estimation results of Table 5) of a one-unit increase in the credit-to-GDP gap on the predicted probability $\Pr(CCyB^{>0} = 1)$ and the corresponding standard errors. The marginal effects are evaluated for the credit-to-GDP gap at 2 percentage points while all other variables are at their means. In line with our previous empirical results, these alternative hypotheses are rejected at conventional significance levels.

Table 6. Additional χ^2-Tests

Pr($CCyB^{>0} = 1$)	**I**		**II**		**III**		**IV**	
	dy/dx	**SE**	dy/dx	**SE**	dy/dx	**SE**	dy/dx	**SE**
Credit-to-GDP Gap	0.0015	0.0041	0.0008	0.0026	0.0017	0.0049	0.0009	0.0027
$H_0 : dy/dx$ **Equals**	χ^2 (1)	**Prob** $> \chi^2$	χ^2 (1)	**Prob** $> \chi^2$	χ^2 (1)	**Prob** $> \chi^2$	χ^2 (1)	**Prob** $> \chi^2$
0	0.13	0.7229	0.09	0.7583	0.12	0.7312	0.10	0.7538
0.25	3636.43	0.0000	8871.58	0.0000	2592.40	0.0000	8254.00	0.0000
0.5	14631.36	0.0000	35602.33	0.0000	10439.66	0.0000	33130.11	0.0000
0.75	32984.94	0.0000	80192.35	0.0000	23541.91	0.0000	74628.42	0.0000
1	58697.15	0.0000	$1.4 * 10^5$	0.0000	41899.15	0.0000	$1.3 * 10^5$	0.0000

Note: The table reports conditional marginal effects dy/dx (based on the estimation results of Table 5) of a one-unit increase in the credit-to-GDP gap on the predicted probability Pr($CCyB^{>0} = 1$) and the corresponding standard errors (Delta-method). The marginal effects are evaluated for the credit-to-GDP gap at 2 percentage points while all other variables are at their means. The 2 percentage points threshold is used, as the ESRB (2014) recommendation assigns positive buffer benchmark values for gaps exceeding 2 percentage points. The test statistics refer to the logistic regression results of Table 5. The χ^2-test tests the null hypothesis that the conditional marginal effect dy/dx of a one-unit increase in the credit-to-GDP gap on the predicted probability Pr($CCyB^{>0} = 1$) equals 0, 25, 50, 75, and 100 percentage points, respectively.

While our results indicate that the credit-to-GDP gap does not determine CCyB policies, this does not mean that the designated authorities decide in a discretionary way only. There are elements of "guidance" present in European CCyB policies. It seems that designated authorities use some of the indicators that have been proposed as more discretionary elements in a rather systematic, almost rule-based manner. Increases in house price growth, e.g., are significantly associated with higher log-odds ratios in the binary CCyB variable (specifications I–IV in Table 5). Stronger house price inflation increases the probability that national designated authorities make use of the CCyB. Quantitatively, a one-standard-deviation increase in house price growth (versus its mean) raises the probability of using the buffer approximately by 6 percent to 8 percent, given all other covariates are at their means. This contrasts with Edge and Liang (2020) who do not find a significant relationship between positive CCyBs and house price changes. As they use annual data, their approach might not be able to pick up the dynamics of house price inflation and subsequent reactions of the regulators.

Also, an increase in distressed credit tends to lower the likelihood of the CCyB requirement, as indicated by a significant and negative coefficient for the non-performing loans as a percentage of total gross loans variable. The negative sign is consistent with the stabilizing objective of the CCyB, namely, to build up buffers under favorable economic conditions when the share of non-performing loans is low. The CCyB provides a preemptive cushion to be built up in good times when accumulating additional capital via retained earnings and raising capital is relatively easy (Couaillier, Idier, and Scalone 2019). In bad times, the CCyB allows the release of capital to support banks in providing sufficient credit to the real economy, even when experiencing unexpected write-offs (ESRB 2014). Given that non-performing loans are included as a contemporaneous variable, rising shares of non-performing loans signal that risks are already materializing to some extent, which implies a reduction of capital requirements as a countercyclical measure. Please note that the share of non-performing loans has been decreasing in almost all countries during the observation period. Interestingly, as with the credit-to-GDP gap, we do not find robust links to other systemic risk indicators mentioned before. This might reflect heterogeneous cross-sectional policy responses (e.g., ESRB 2019) when taking additional risk indicators into account.

Institutional indicator variables that reflect which specific policymaker is ultimately responsible for CCyB decisions are generally not significant. However, coefficients of the prudential regulator and the central bank are always negative. When controlling only for the Financial Stability Committee as the decisionmaking authority (FSC sets CCyB), the coefficient is positive, however, at an insignificant level. In contrast, the coefficient was significant in our robustness exercises. Overall, the results support the findings of Edge and Liang (2020). The probability of a positive CCyB is lower if the central bank or the prudential regulator decides. For the prudential regulator, the reduced likelihood to activate the countercyclical buffer may be explained by the focus—and possibly preference—on microprudential policy (Edge and Liang 2020). Countries use the CCyB more likely if the FSC takes the final decision. FSCs that can set the CCyB directly are relatively powerful. Given their macroprudential focus, it is not surprising that they use the capital buffer more often.

So far, we have examined whether or not the countercyclical capital buffer is used, regardless of the specific setting of the rate. This aspect is in particular relevant for the decision of designated authorities to use the CCyB at all. In a second step, we analyze a complementary question, namely how decisionmakers vary CCyB rates with respect to the macrofinancial environment by estimating the following linear unobserved effects model

$$CCyB_{it} = \alpha + \mathbf{x}'_{it}\boldsymbol{\beta} + u_i + v_t + \epsilon_{it}, \tag{3}$$

where $CCyB_{it}$ denotes the latest pending rate of the countercyclical capital buffer in country i for quarter t, α a constant, $\mathbf{x}_{it}$ the vector of aforementioned risk indicators for country i in quarter t, and $\boldsymbol{\beta}$ the corresponding parameters. u_i is the unobserved country effect, v_t the aggregate time effect, and ϵ_{it} the error term. We only include observations of countries that have already announced non-zero CCyB rates at at least one point in time through 2019.

Table 7 summarizes the results of the linear (fixed-effects) regression with the announced buffer rate as the dependent variable. Again, as the insignificant coefficients of the credit-to-GDP gap

Table 7. Linear Regression

CCyB	**I**	**II**	**III**
Credit-to-GDP Gap	0.0107	0.0084	0.0021
	(0.0071)	(0.0058)	(0.0041)
Credit Growth (1Y)	0.0156	0.0124	0.0088
	(0.0131)	(0.0133)	(0.0059)
House Prices (5Y)	0.0290**	0.0190*	0.0092
	(0.0096)	(0.0099)	(0.0073)
Stock Index (1Y)	–0.0029	–0.0044	0.0007
	(0.0029)	(0.0030)	(0.0035)
Stock Index Volatility	–0.0497	0.0182	0.0226
	(0.1614)	(0.0953)	(0.0521)
Current Account	–0.0082	–0.0014	0.0001
	(0.0072)	(0.0031)	(0.0026)
Regulatory Capital	0.0035	0.0500	0.0278
	(0.0654)	(0.0421)	(0.0211)
Non-performing Loans	–0.0069	–0.1171**	–0.0337
	(0.0422)	(0.0423)	(0.0288)
Country FE	No	Yes	Yes
Year Effects	No	No	Yes
Observations	229	229	229
$R^2(within)$	0.42	0.52	0.72

Note: The dependent variable is the announced CCyB level. All observations of countries were used which have announced positive CCyB rates at least once within the observation period. The models were estimated using a constant, which is not reported. Clustered standard errors (at country level) are reported in parentheses. Significance levels: ***1 percent, **5 percent, *10 percent. Data sources are provided in Table 1.

(specifications I–III in Table 7) indicate, there is no evidence that designated authorities base their CCyB decisions systematically on the officially recommended credit-to-GDP gap.

Analogous to the above discussion, we investigate in greater detail the potential role of the credit-to-GDP gap as the recommended rule-based element in CCyB policies. For the coefficients of the credit-to-GDP gap in Table 7 we perform additional F-tests (reported in Table 8). The coefficients are not tested against zero but against the linear slope parameter of the recommended

Table 8. Additional F-Tests

Table	Coefficient	H_0	I		II		III	
			F Statistic	Prob > F	F Statistic	Prob > F	F Statistic	Prob > F
7	Credit-to-GDP Gap	$\beta = 0.3125$	F(1,11) = 1831.79	0.0000	F(1,11) = 2722.99	0.0000	F(1,11) = 5854.55	0.0000
10	Credit-to-GDP Gap	$\beta = 0.3125$	F(1,11) = 1083.30	0.0000	F(1,11) = 1493.49	0.0000	F(1,11) = 3748.66	0.0000
12	Credit-to-GDP Gap	$\beta = 0.3125$	F(1,9) = 1620.66	0.0000	F(1,9) = 4042.28	0.0000	F(1,9) = 4327.57	0.0000
14	Credit-to-GDP Gap	$\beta = 0.3125$	F(1,11) = 1850.34	0.0000	F(1,11) = 3350.67	0.0000	F(1,11) = 10463.67	0.0000
17	Credit-to-GDP Gap	$\beta = 0.3125$	F(1,11) = 2922.61	0.0000	F(1,11) = 1113.29	0.0000	F(1,11) = 5069.77	0.0000
17	Buffer Guide	$\beta = 1$	F(1,11) = 9.70	0.0098	F(1,11) = 35.54	0.0001	F(1,11) = 102.57	0.0000

Note: The F-test tests the null hypothesis $\beta = 0.3125$ for the credit-to-GDP gap and $\beta = 1$ for the buffer guide, respectively.

buffer benchmark rule, H_0: $\beta_{Credit-to-GDP\ gap} = 0.3125$. The results of these tests imply that the alternative null hypotheses are clearly rejected on conventional significance levels, i.e., authorities do not set the CCyB according to the rule-based component (Equation (1)).

Once again, house price growth seems to be policy relevant, at least if we do not account for aggregate time effects (specification I–II). Thus, higher house price growth is associated not only with an increasing probability of setting a positive CCyB but also with higher rates, given the buffer is already activated. Quantitatively, an increase in house price inflation by 10 percentage points is associated with a rise in the buffer of approximately 0.2–0.3 percentage point. When we control for time effects, the sign remains robust while the coefficient becomes insignificant. This pattern may be caused by time trends, which are captured by the aggregate time effects in specification III.

When we control for both country fixed effects and aggregate time effects (specification III), all explanatory variables become insignificant. We will further elaborate on this finding in our robustness exercises.

In contrast to the binary response regression, improved credit quality, as measured by a decreased non-performing loans ratio, does not seem to result consistently in significantly higher buffer rates, given that the country has already implemented a CCyB policy. Country and time effects seem to play a crucial role when considering domestic non-performing loans ratios. This result might reflect that the standard deviation of non-performing loans within a given country is much lower than the standard deviation between different countries.

Taken together, our empirical results indicate that the policy to set a specific buffer rate should be distinguished from the general decision on using the CCyB at all. However, and not surprisingly, there exists considerable overlap. For both decisions, we do not find robust evidence that the credit-to-GDP gap is relevant—despite its prominent role in official communications. In contrast, in both decisions, house price inflation seems to play an important, systematic role. In the case of other risk indicators listed in the ESRB (2014) recommendation, policymakers do not seem to focus on external imbalances or—interestingly—current regulatory bank capital. For the role of equity prices and their role in the buildup of risk, the timing might be crucial.[7]

5. Robustness

In the subsequent robustness analysis, we account for additional variables that signal the buildup of risk and discuss data availability issues as well as alternative estimation approaches.

National authorities are required to announce the credit-to-GDP ratio each quarter together with the credit-to-GDP gap (Directive 2013/36/EU 2013, Article 136). To account for diverging signals of alternative credit measures, we include the credit-to-GDP ratio[8] alongside the credit-to-GDP gap in the random-effects logit model and in the linear panel model. The ratio provides a debt measure standardized by the country's GDP.

As reported in Table 9, the main findings in specifications I–IV do not alter. The coefficients of house price inflation are positive on the 1 percent significance level and have similar magnitudes. The negative effect of the contemporaneous share of non-performing loans is also robust against the additional consideration of the credit-to-GDP ratio. The institutional indicator variables have the expected sign, i.e., negative coefficients for the prudential regulator and the central bank and a positive coefficient for the FSC. In terms of the linear level regression (Table 10), credit gaps remain insignificant. The coefficients of house price growth are qualitatively unaffected by the consideration of the additional variable. In accordance with our previous results (Table 7), house price growth seems to be less critical when controlling both for country fixed effects and aggregate time effects.

Given the fundamental role of bank-based financing in Europe, decisionmakers may focus more on bank credit than total non-financial debt. To assess if the specific measure of credit is crucial for our findings, we replace broad credit with bank credit (loans

[7]For instance, Borio and Lowe (2002a) found that equity price gaps peak earlier than other risk indicators.

[8]We retrieved credit-to-GDP ratios from the ESRB and national authorities. In most cases, we used the ratio based on the broad credit aggregate. However, for some countries, the ratio is available based on narrower aggregates only.

Table 9. Random-Effects Logistic Regression (robustness—including credit-to-GDP ratio)

$CCyB^{>0}$	I	II	III	IV
Credit-to-GDP Gap	0.0724	0.0966	0.0908	0.0782
	(0.0594)	(0.0594)	(0.0607)	(0.0580)
Credit Growth (1Y)	–0.2293	–0.2235	–0.2354	–0.2206
	(0.1531)	(0.1532)	(0.1541)	(0.1512)
Credit-to-GDP Ratio	0.0237	0.0161	0.0267	0.0120
	(0.0248)	(0.0256)	(0.0251)	(0.0230)
House Prices (5Y)	0.3406***	0.3513***	0.3395***	0.3427***
	(0.0482)	(0.0465)	(0.0447)	(0.0448)
Stock Index (1Y)	–0.0460	–0.0453	–0.0469	–0.0462
	(0.0353)	(0.0350)	(0.0358)	(0.0346)
Stock Index Volatility	0.9340	1.1609	1.0572	0.9966
	(1.2959)	(1.3077)	(1.3102)	(1.2788)
Current Account	0.0050	0.0069	0.0060	0.0037
	(0.0565)	(0.0572)	(0.0578)	(0.0549)
Regulatory Capital	0.1278	0.0717	0.1818	0.1018
	(0.2723)	(0.2676)	(0.2709)	(0.2640)
Non-performing Loans	–2.8586***	–2.8453***	–2.9178***	–2.7471***
	(0.5067)	(0.5189)	(0.5366)	(0.4772)
PR Sets CCyB		–13.2799	–7.9125**	
		(9.6087)	(3.8240)	
CB Sets CCyB		–5.3819		
		(9.3447)		
MF Sets CCyB		–7.1296		
		(10.7378)		
FSC Sets CCyB				8.9232*
				(4.7253)
Observations	493	493	493	493
Log-Likelihood	–77.82	–77.36	–77.38	–77.61
χ^2 (DF)	109.34 (9)	132.43 (12)	115.13 (10)	140.84 (10)

Note: The dependent variable is the binary CCyB decision. The models were estimated using a constant, which is not reported. Significance levels: ***1 percent, **5 percent, *10 percent. Data sources are provided in Table 1.

and debt securities granted by monetary financial institutions). As data are missing for some countries, the number of observations shrinks slightly. There are no material differences in the binary case (Table 11), as house prices and non-performing loans seem to be the main drivers of CCyB policies. However, we do not see a clear pattern for adjustments over time (Table 12), as the house price

Table 10. Linear Regression (robustness—including credit-to-GDP ratio)

CCyB	I	II	III
Credit-to-GDP Gap	0.0134	0.0042	0.0030
	(0.0091)	(0.0080)	(0.0051)
Credit Growth (1Y)	0.0188	0.0141	0.0083
	(0.0119)	(0.0133)	(0.0060)
Credit-to-GDP Ratio	0.0019	0.0056	–0.0012
	(0.0024)	(0.0058)	(0.0033)
House Prices (5Y)	0.0298**	0.0173*	0.0094
	(0.0104)	(0.0092)	(0.0073)
Stock Index (1Y)	–0.0039*	–0.0044	0.0007
	(0.0020)	(0.0027)	(0.0036)
Stock Index Volatility	–0.1251	0.0510	0.0167
	(0.1761)	(0.0799)	(0.0567)
Current Account	–0.0086	–0.0008	–0.0000
	(0.0070)	(0.0030)	(0.0026)
Regulatory Capital	–0.0080	0.0528	0.0268
	(0.0649)	(0.0425)	(0.0214)
Non-performing Loans	–0.0020	–0.1304***	–0.0294
	(0.0511)	(0.0417)	(0.0293)
Country FE	No	Yes	Yes
Year Effects	No	No	Yes
Observations	229	229	229
$R^2(within)$	0.43	0.53	0.72

Note: The dependent variable is the announced CCyB level. All observations of countries were used which have announced positive CCyB rates at least once within the observation period. The models were estimated using a constant, which is not reported. Clustered standard errors (at country level) are reported in parentheses. Significance levels: ***1 percent, **5 percent, *10 percent. Data sources are provided in Table 1.

variable becomes insignificant when we control for the unobserved country and time effects.

While the credit-to-GDP ratio and gap data are those available at the time of decision (ESRB data set), we typically use contemporary observations that have not been publicly available at the time of decision in the case of the additional variables. Thus, we implicitly assume that decisionmakers have a considerable information advantage for these variables. As a further problem, we only have ex post revised time series that might differ from those available at the time

Table 11. Random-Effects Logistic Regression (robustness—MFI credit growth)

$CCyB^{>0}$	I	II	III
Credit-to-GDP Gap	0.0130	0.0064	0.0045
	(0.0658)	(0.0697)	(0.0756)
MFI Credit Growth (1Y)	0.2463	0.1974	0.1977
	(0.1912)	(0.1920)	(0.2374)
House Prices (5Y)	0.3918***	0.3753***	0.4209***
	(0.0466)	(0.0495)	(0.0855)
Stock Index (1Y)	–0.0154	–0.0156	–0.0140
	(0.0366)	(0.0355)	(0.0392)
Stock Index Volatility	2.4010	2.1623	2.5014
	(1.6737)	(1.6434)	(1.8624)
Current Account	–0.0078	–0.0051	–0.0036
	(0.0522)	(0.0525)	(0.0557)
Regulatory Capital	–0.0928	–0.1261	0.0260
	(0.3043)	(0.2943)	(0.3362)
Non-performing Loans	–2.5832***	–2.5446***	–2.8701***
	(0.4465)	(0.4813)	(0.5586)
PR Sets CCyB		–1.0146	–3.2881
		(6.4524)	(4.3199)
CB Sets CCyB		–1.3422	
		(5.4938)	
Observations	454	454	454
Log-Likelihood	–60.20	–60.50	–60.29
χ^2 (DF)	149.83	119.68	51.86

Note: The dependent variable is the binary CCyB decision. The models were estimated using a constant, which is not reported. Significance levels: ***1 percent, **5 percent, *10 percent. Data sources are provided in Table 1.

of the decision. To account for potential information lags, we regress the CCyB on the first lags of the independent variables other than credit-to-GDP data, stock market variables, and the institutional indicators. Credit-to-GDP gaps and ratios from the ESRB data set were available at the time of the decision. Hence, we do not have to account for further information lags. Similarly, we do not use lagged values of the stock market data, as stock index data are available in real time. We notice differences for the CCyB indicator regression (Table 13) as we identify more significant coefficients. While the influence of house price inflation and non-performing loans does not change qualitatively, the credit-to-GDP gap and stock price changes

Table 12. Linear Regression (robustness—MFI credit growth)

CCyB	I	II	III
Credit-to-GDP Gap	0.0123	0.0058	0.0038
	(0.0075)	(0.0048)	(0.0047)
MFI Credit Growth (1Y)	–0.0303	–0.0390	–0.0177
	(0.0378)	(0.0214)	(0.0209)
House Prices (5Y)	0.0304**	0.0207*	0.0090
	(0.0119)	(0.0097)	(0.0074)
Stock Index (1Y)	–0.0008	–0.0020	0.0022
	(0.0030)	(0.0032)	(0.0038)
Stock Index Volatility	–0.0826	0.0263	0.0249
	(0.1837)	(0.1039)	(0.0627)
Current Account	–0.0117	–0.0023	–0.0000
	(0.0067)	(0.0028)	(0.0024)
Regulatory Capital	–0.0159	0.0265	0.0340
	(0.0693)	(0.0431)	(0.0310)
Non-performing Loans	–0.0194	–0.1334**	–0.0631*
	(0.0515)	(0.0443)	(0.0294)
Country FE	No	Yes	Yes
Year Effects	No	No	Yes
Observations	190	190	190
$R^2(within)$	0.42	0.56	0.73

Note: The dependent variable is the announced CCyB level. All observations of countries were used which have announced positive CCyB rates at least once within the observation period. The models were estimated using a constant, which is not reported. Clustered standard errors (at country level) are reported in parentheses. Significance levels: ***1 percent, **5 percent, *10 percent. Data sources are provided in Table 1.

become important. Consistent with intuition and the ESRB (2014) recommendation, higher credit gaps are associated with a higher likelihood to use the capital buffer. The negative sign of the year-on-year change of stock prices is against intuition, which states that higher equity valuations may indicate a buildup of systemic risk. Interestingly, we also identify more significant coefficients for the linear level regression. Credit growth and house price inflation are relevant when we control for the unobserved country and aggregate time effects (Table 14, specification III). The coefficient of the credit-to-GDP gap is insignificant in all of the three specifications.

Analogous to the approach of Edge and Liang (2020), we check if the selection of the specific logit models is crucial for our outcomes.

Table 13. Random-Effects Logistic Regression (lags)

$CCyB^{>0}$	I	II	III	IV
Credit-to-GDP Gap	0.1029	0.1061*	0.1158*	0.0872**
	(0.0672)	(0.0627)	(0.0644)	(0.0391)
Credit Growth (1Y, L1)	–0.2171	–0.2300	–0.2343	–0.1327
	(0.1510)	(0.1475)	(0.1525)	(0.1066)
House Prices (5Y, L1)	0.3802***	0.3530***	0.3738***	0.2471***
	(0.0506)	(0.0453)	(0.0464)	(0.0322)
Stock Index (1Y)	–0.0962***	–0.0963***	–0.0983***	–0.0786***
	(0.0364)	(0.0354)	(0.0362)	(0.0249)
Stock Index Volatility	–1.1079	–1.0599	–1.1061	–0.9390
	(0.8720)	(0.8684)	(0.8750)	(0.7498)
Current Account (L1)	0.0009	0.0083	0.0079	–0.0080
	(0.0545)	(0.0548)	(0.0564)	(0.0418)
Regulatory Capital (L1)	0.3469	0.4245*	0.4099*	0.2001
	(0.2553)	(0.2257)	(0.2351)	(0.1609)
Non-performing Loans (L1)	–2.6573***	–2.6122***	–2.7503***	–1.6300***
	(0.4876)	(0.4673)	(0.4805)	(0.2616)
PR Sets CCyB		–14.7971	–8.8388**	
		(9.6831)	(4.0135)	
CB Sets CCyB		–7.3223		
		(9.0534)		
MF Sets CCyB		–7.4182		
		(10.3989)		
FSC Sets CCyB				3.4595*
				(1.8554)
Observations	512	512	512	512
Log-Likelihood	–78.62	–77.71	–77.84	–83.10
χ^2 (DF)	144.46 (8)	155.61 (11)	156.01 (9)	257.47 (9)

Note: The dependent variable is the binary CCyB decision. The models were estimated using a constant, which is not reported. Significance levels: ***1 percent, **5 percent, *10 percent. Data sources are provided in Table 1.

Amemiya (1981) argues that probit and logit models lead to similar results as long as the data are not strongly concentrated at the end of the probability distribution. This concentration could be an issue since many non-positive CCyBs imply a high distribution mass at zero. Amemiya (1981) shows that logit coefficients can be approximately converted into probit estimates by applying the formula $\hat{\beta}_L = 1.6\hat{\beta}_P$, where $\hat{\beta}_L$ denotes the logit coefficient and $\hat{\beta}_P$ the probit estimate. Again, the credit-to-GDP gap is insignificant in the probit estimation (Table 15). In contrast, house prices remain significant. The probit estimates show stronger evidence for the negative

Table 14. Linear Regression (lags)

CCyB	**I**	**II**	**III**
Credit-to-GDP Gap	0.0125	0.0076	0.0042
	(0.0070)	(0.0053)	(0.0030)
Credit Growth (1Y, L1)	0.0141	0.0182*	0.0128**
	(0.0115)	(0.0096)	(0.0053)
House Prices (5Y, L1)	0.0315***	0.0227**	0.0174**
	(0.0089)	(0.0083)	(0.0071)
Stock Index (1Y)	–0.0026	–0.0054*	0.0005
	(0.0027)	(0.0029)	(0.0036)
Stock Index Volatility	–0.0980	–0.1438	–0.0271
	(0.1353)	(0.0893)	(0.0578)
Current Account (L1)	–0.0062	–0.0003	–0.0007
	(0.0080)	(0.0018)	(0.0027)
Regulatory Capital (L1)	–0.0041	0.0403	0.0286
	(0.0636)	(0.0334)	(0.0227)
Non-performing Loans (L1)	0.0112	–0.0740*	–0.0250
	(0.0413)	(0.0410)	(0.0336)
Country FE	No	Yes	Yes
Year Effects	No	No	Yes
Observations	236	236	236
R^2 (*within*)	0.45	0.55	0.69

Note: The dependent variable is the announced CCyB level. All observations of countries were used which have announced positive CCyB rates at least once within the observation period. The models were estimated using a constant, which is not reported. Clustered standard errors (at country level) are reported in parentheses. Significance levels: ***1 percent, **5 percent, *10 percent. Data sources are provided in Table 1.

(positive) impact on the likelihood of setting positive CCyBs when the prudential regulator (the financial stability committee) decides.

As the independent variables measure systemic risk in different dimensions, it should be informative to inspect the co-movements of the explanatory variables when interpreting multivariate regression results.[9] Coefficients of correlation are reported for the continuous explanatory variables (Table 3). None of the bivariate correlations exceeds 0.6. In our baseline regressions (Table 5 and Table 7) in which we do not include the MFI credit growth, the highest bivariate correlation is below 0.5. We also calculated the centered variance

[9]We thank an anonymous referee for pointing out this aspect.

Table 15. Random-Effects Probit Regression

$CCyB^{>0}$	I	II	III	IV
Credit-to-GDP Gap	0.0393 (0.0336)	0.0495 (0.0340)	0.0446 (0.0342)	0.0507 (0.0368)
Credit Growth (1Y)	–0.1222 (0.0842)	–0.0928 (0.0735)	–0.1171 (0.0817)	–0.1368 (0.0921)
House Prices (5Y)	0.2136*** (0.0277)	0.1762*** (0.0214)	0.2011*** (0.0331)	0.2294*** (0.0313)
Stock Index (1Y)	–0.0258 (0.0200)	–0.0256 (0.0177)	–0.0267 (0.0197)	–0.0231 (0.0206)
Stock Index Volatility	0.6389 (0.7372)	0.4689 (0.6626)	0.6550 (0.7367)	0.8075 (0.7723)
Current Account	–0.0040 (0.0302)	–0.0068 (0.0265)	–0.0029 (0.0293)	–0.0015 (0.0317)
Regulatory Capital	0.1189 (0.1537)	0.1415 (0.1423)	0.1537 (0.1575)	0.1220 (0.1509)
Non-performing Loans	–1.5564*** (0.2190)	–1.2757*** (0.1808)	–1.4997*** (0.2000)	–1.7006*** (0.2960)
PR Sets CCyB		–5.5355*** (1.4388)	–4.1937* (2.1658)	
CB Sets CCyB		–2.9510* (1.6157)		
MF Sets CCyB		–2.2030 (2.6578)		
FSC Sets CCyB				5.2927*** (2.0268)
Observations	493	493	493	493
Log-Likelihood	–78.09	–77.41	–77.92	–77.12
χ^2 (DF)	124.48 (8)	169.89 (11)	109.06 (9)	82.31 (9)

Note: The dependent variable is the binary CCyB decision. The models were estimated using a constant, which is not reported. Significance levels: ***1 percent, **5 percent, *10 percent. Data sources are provided in Table 1.

inflation factors (VIFs) for all continuous variables included in the linear-level regression without unobserved effects (column 1 in the linear regression, Table 7). The resulting VIFs (not reported) are small and, as indicated by the bivariate correlation measures, do not show a severe multicollinearity problem.

We also performed univariate analyses by regressing the CCyB level and the binary CCyB indicator variable on all continuous variables separately.

Table 16. Binary Regression—Univariate

Dependent Variable	Independent Variable	Coefficient
$CCyB^{>0}$	Buffer Guide	–0.3611 (0.5373)
$CCyB^{>0}$	Credit-to-GDP Gap	–0.0105 (0.0194)
$CCyB^{>0}$	Credit Growth (1Y)	0.0306 (0.0354)
$CCyB^{>0}$	House Prices (5Y)	0.2622*** (0.0510)
$CCyB^{>0}$	Stock Index (1Y)	–0.0468*** (0.0130)
$CCyB^{>0}$	Stock Index Volatility	–1.0410** (0.5128)
$CCyB^{>0}$	Current Account	0.0050 (0.0220)
$CCyB^{>0}$	Regulatory Capital	0.1262 (0.1062)
$CCyB^{>0}$	Non-performing Loans	–4.4152*** (0.8738)
Observations		493

Note: The dependent variable is the binary CCyB decision. The models were estimated using a constant, which is not reported. Significance levels: ***1 percent, **5 percent, *10 percent. Data sources are provided in Table 1.

Intuitively, if the "rules-based approach" was the main driver for CCyB decisions, we would expect a strong positive and significant relationship between the buffer guide and the CCyB. It may be helpful here to think of an "ideal world" in which the credit-to-GDP gap—and hence the buffer benchmark—is a measure accepted by all national designated authorities that properly reflects the risks in the financial sector. We, therefore, regressed CCyB decisions on the buffer guide (derived from the credit-to-GDP gap), which we took from the ESRB data set. For the linear case, the slope parameter should be approximately equal to one. As shown in Table 16 and Table 17, the buffer guide was neither significantly different from zero for the CCyB indicator variable nor the buffer level. Consistent with our previous regressions, we considered only countries that have used the macroprudential instrument at least once within the observation period in the latter case. The null hypothesis that the linear

Table 17. Linear Regression—Univariate

Dependent Variable	Independent Variable	I	II	III
CCyB	Buffer Guide	0.1669	–0.3945	–0.1702
		(0.2675)	(0.2339)	(0.1155)
CCyB	Credit-to-GDP Gap	0.0055	–0.0090	–0.0021
		(0.0057)	(0.0096)	(0.0044)
CCyB	Credit Growth (1Y)	0.0231	0.0068	0.0066
		(0.0194)	(0.0098)	(0.0041)
CCyB	House Prices (5Y)	0.0243**	0.0296***	0.0098
		(0.0106)	(0.0084)	(0.0071)
CCyB	Stock Index (1Y)	–0.0057	–0.0079	0.0003
		(0.0035)	(0.0049)	(0.0038)
CCyB	Stock Index Volatility	–0.1354	–0.3853**	–0.0460
		(0.1487)	(0.1382)	(0.0485)
CCyB	Current Account	–0.0010	0.0006	0.0010
		(0.0075)	(0.0034)	(0.0024)
CCyB	Regulatory Capital	0.0502	0.0515	0.0168
		(0.0665)	(0.0670)	(0.0172)
CCyB	Non-performing Loans	–0.0651*	–0.1851***	–0.0421*
		(0.0351)	(0.0566)	(0.0213)
Country FE		No	Yes	Yes
Year Effects		No	No	Yes
Observations		229	229	229

Note: The dependent variable is the announced CCyB level. All observations of countries were used which have announced positive CCyB rates at least once within the observation period. The models were estimated using a constant, which is not reported. Clustered standard errors (at country level) are reported in parentheses. Significance levels: ***1 percent, **5 percent, *10 percent. Data sources are provided in Table 1.

coefficient of the credit-to-GDP gap equals 0.3125 and that of the buffer guide equals 1 is clearly rejected on conventional levels (Table 8). The univariate results in Table 16 support the positive impact of house price inflation and the negative influence of non-performing loans on the likelihood of using the buffer. Both variables also seem relevant for buffer calibration (Table 17), at least when we do not control for the country and time effects. In univariate approaches, the stock market variables seem to be more relevant.

6. Conclusions

Based on its "guided discretion" approach, the European Systemic Risk Board recommends a prominent role of the credit-to-GDP gap and the related benchmark buffer rate. However, our empirical analysis indicates that the credit-to-GDP gap, as the rule-based element, seems to be only of a minor, if any, relevance for national macroprudential policies.

Interestingly, that does not mean that national authorities act in a downright discretionary way only. We find that policymakers systematically take into account some of the other risk indicators related to the financial cycle as suggested by the ESRB (2014). In particular, they seem to react to house price inflation when setting the countercyclical buffer rate for domestic exposure. This is likely to reflect concerns about potential overvaluations in real estate markets, the subsequent risk of bursting housing bubbles, and distress in the banking sector. As pointed out by Borio and Lowe (2002b) and Borio (2014) among others, real estate prices are a key driver of the financial cycle. Also, credit quality, as measured by the non-performing loans ratio, appears to play an important role in setting the countercyclical capital buffer.

Our empirical results are related to a conflict that has been discussed at great length in the field of monetary policy. In choosing their policy framework, policymakers not only have the choice between (pure) rules versus (pure) discretion; they can also choose to constrain discretion by implementing *rule-like* features (Mishkin 2017). A similar logic might hold in the field of macroprudential policy. By strengthening rule-like elements in their policy decisions, authorities could possibly improve the efficiency of their policies. Rule-based components enhance the comparability of macroprudential policy among different countries and should make decisions more comprehensible to financial markets. Transparent communication of indicators and their consistent application could improve the predictability of capital buffer decisions and reduce adaption costs for financial institutions.

Unsurprisingly, some caveats should be kept in mind. Since the CCyB is a relatively novel instrument, our analysis does not cover policy decisions for the entire financial cycle. Moreover, the reporting of consistent data (e.g., credit-to-GDP gaps, credit-to-GDP ratios) on the European level is still in its infancy. Consistently calculated and published indicators would help to improve the analysis of European CCyB decisions.

Finally, it is puzzling that national authorities do not stick more closely to the buffer guide rule they have agreed to as ESRB members. Apparently, they are not at odds with a systematic CCyB policy, at least when it is based on indicators such as house price inflation and credit quality. By not following the officially agreed-upon credit-to-GDP rule while concentrating on complementary variables, they pursue rather non-transparent and inconsistent policies. They neglect the potentially relevant information channel of their policies and forgo the benefits of a time-consistent policy. In this situation, the following two options seem available. Either the ESRB recommendations are brought in line with the current CCyB policies on the national level, or else national buffer decisions should be more closely linked to the single quantitative rule.

References

Akinci, O., and J. Olmstead-Rumsey. 2018. "How Effective Are Macroprudential Policies? An Empirical Investigation." *Journal of Financial Intermediation* 33 (January): 33–57.

Amemiya, T. 1981. "Qualitative Response Models: A Survey." *Journal of Economic Literature* 19 (4): 1483–1536.

Babić, D., and S. Fahr. 2019. "Shelter from the Storm: Recent Countercyclical Capital Buffer (CCyB) Decisions." *Macroprudential Bulletin* (European Central Bank) 7 (March 27).

BaFin. 2019. "General Administrative Act Governing the Rate for the Domestic Countercyclical Capital Buffer under Section 10d of the KWG." https://www.bafin.de/SharedDocs/Veroeffentlichungen/EN/Aufsichtsrecht/Verfuegung/vf_190628_allgvfg_antizykl_kapitalpuffer_en.html. Last checked on June 17, 2021.

Banca d'Italia. 2019. "The Countercyclical Capital Buffer (CCyB) Rate for the Fourth Quarter of 2019 Has Been Set at Zero Per Cent." https://www.bancaditalia.it/compiti/stabilita-finanziaria/politica-macroprudenziale/ccyb-4-2019/index.html. Last checked on June 17, 2021.

Banco de Portugal. 2019. "Countercyclical Capital Buffer — Background Document for the 3rd Quarter 2019." https://www.bportugal.pt/sites/default/files/quarter_3_2019_docccb_en.pdf. Last checked on June 17, 2021.

Bank for International Settlements. 2017. "Implementation — Range of Practices in Implementing the Countercyclical Capital Buffer." https://www.bis.org/bcbs/publ/d407.pdf. Last checked on August 25, 2020.

Basel Committee on Banking Supervision. 2010. "Guidance for National Authorities Operating the Countercyclical Capital Buffer." Bank for International Settlements. https://www.bis.org/publ/bcbs187.pdf. Last checked on September 9, 2020.

Borio, C. 2014. "The Financial Cycle and Macroeconomics: What Have We Learnt?" *Journal of Banking and Finance* 45 (August): 182–98.

Borio, C., M. Drehmann, L. Gambacorta, G. Jiménez, and C. Trucharte. 2010. "Countercyclical Capital Buffers: Exploring Options." BIS Working Paper No. 317.

Borio, C., and P. Lowe. 2002a. "Assessing the Risk of Banking Crises." *BIS Quarterly Review* 7 (1): 43–54.

———. 2002b. "Asset Prices, Financial and Monetary Stability: Exploring the Nexus." BIS Working Paper No. 114.

Brunnermeier, M. K. 2009. "Deciphering the Liquidity and Credit Crunch 2007-2008." *Journal of Economic Perspectives* 23 (1): 77–100.

Cerutti, E., S. Claessens, and L. Laeven. 2017. "The Use and Effectiveness of Macroprudential Policies: New Evidence." *Journal of Financial Stability* 28 (February): 203–24.

Cerutti, E., R. Correa, E. Fiorentino, and E. Segalla. 2016. "Changes in Prudential Policy Instruments—A New Cross-Country Database." IMF Working Paper No. 16/110.

Christiansen, C., M. Schmeling, and A. Schrimpf. 2012. "A Comprehensive Look at Financial Volatility Prediction by Economic Variables." *Journal of Applied Econometrics* 27 (6): 956–77.

Couaillier, C., J. Idier, and V. Scalone. 2019. "Activation of Countercyclical Capital Buffers in Europe: Initial Experiences." *Banque de France Bulletin* 222 (March-April): 1–10.

Czech National Bank. 2015. "Provision of a General Nature IV/2015 on Setting the Countercyclical Capital Buffer Rate for the Czech Republic No. IV/2015." https://www.cnb.cz/en/financial-stability/macroprudential-policy/the-countercyclical-capital-buffer/provision-of-a-general-nature-on-setting-the-count ercyclical-capital-buffer-rate/Provision-of-a-general-nature-IV-2015. Last checked on June 17, 2021.

Detken, C., O. Weeken, L. Alessi, D. Bonfim, M. M. Boucinha, C. Castro, S. Frontczak et al. 2014. "Operationalising the Countercyclical Capital Buffer: Indicator Selection, Threshold Identification and Calibration Options." ESRB Occasional Paper No. 5.

Directive 2013/36/EU. 2013. "DIRECTIVE 2013/36/EU OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 26 June 2013 on Access to the Activity of Credit Institutions and the Prudential Supervision of Credit Institutions and Investment Firms, Amending Directive 2002/87/EC and Repealing Directives 2006/48/EC and 2006/49/EC." *Official Journal of the European Union* 176: 338–436.

Drehmann, M. 2013. "Total Credit as an Early Warning Indicator for Systemic Banking Crises." *BIS Quarterly Review* (June): 41–46.

Drehmann, M., and J. Yetman. 2018. "Why You Should Use the Hodrick-Prescott Filter–at Least to Generate Credit Gaps." BIS Working Paper No. 744.

Edge, R. M., and J. N. Liang. 2020. "Financial Stability Committees and the Countercyclical Capital Buffer." Discussion Paper No. 04/2020, Deutsche Bundesbank.

Edge, R. M., and R. R. Meisenzahl. 2011. "The Unreliability of Credit-to-GDP Ratio Gaps in Real-Time: Implications for Countercyclical Capital Buffers." *International Journal of Central Banking* 7 (4): 261–98.

European Systemic Risk Board. 2014. "Recommendation of the European Systemic Risk Board of 18 June 2014 on Guidance for Setting Countercyclical Buffer Rates (ESRB/2014/1)." https://www.esrb.europa.eu/pub/pdf/recommendations/140630_ESRB_Recommendation.en.pdf?90ace214e21806d5347f5b45d4fc541a. Last checked on August 25, 2020.

———. 2019. "A Review of Macroprudential Policy in the EU in 2018." https://www.esrb.europa.eu/pub/pdf/reports/esrb~32aae4bd95.report190430_reviewofmacroprudentialpolicy.pdf. Last checked on August 25, 2020.

———. 2020. "A Review of Macroprudential Policy in the EU in 2019." https://www.esrb.europa.eu/pub/pdf/reports/review_macroprudential_policy/esrb.report200429_reviewofmacr oprudentialpolicy~13aab65584.en.pdf. Last checked on August 25, 2020.

Finansinspektionen. 2018. "Decision Memorandum: Amendment to Regulations Regarding the Countercyclical Buffer Rate, FI Ref.18-11833." https://www.fi.se/globalassets/media/dokument/fffs-bilagor/2018/beslutspm-fffs-2018-17_eng.pdf. Last checked on June 17, 2021.

Galán, J. 2019. "Measuring Credit-to-GDP Gaps. The Hodrick-Prescott Filter Revisited." Occasional Paper No. 1906, Banco de Espana.

Gischer, H., B. Herz, and L. Menkhoff. 2019. "Anti-cyclical Capital Buffer Activated — Too Little, Too Late and Still Right." *Wirtschaftsdienst* 99 (11): 784–88.

Hamilton, J. D. 2018. "Why You Should Never Use the Hodrick-Prescott Filter." *Review of Economics and Statistics* 100 (5): 831–43.

Kauko, K., and E. Tölö. 2019. "On the Long-Run Calibration of the Credit-to-GDP Gap as a Banking Crisis Predictor." Research Discussion Paper No. 6/2019, Bank of Finland.

Mishkin, F. S. 2017. "Making Discretion in Monetary Policy More Rule-like." NBER Working Paper No. 24135.

Reigl, N., and L. Uusküla. 2018. "Alternative Frameworks for Measuring Credit Gaps and Setting Countercyclical Capital Buffers." Working Paper No. 7/2018, Eesti Pank.

Tente, N., I. Stein, L. Silbermann, and T. Deckers. 2015. "The Countercyclical Capital Buffer in Germany — Analytical Framework for the Assessment of an Appropriate Domestic Buffer Rate." Deutsche Bundesbank, November.

Tölö, E., H. Laakkonen, and S. Kalatie. 2018. "Evaluating Indicators for Use in Setting the Countercyclical Capital Buffer." *International Journal of Central Banking* 14 (2): 51–112.

Wezel, T. 2019. "Conceptual Issues in Calibrating the Basel III Countercyclical Capital Buffer." IMF Working Paper No. 19/86.

Wolf, E., F. Mokinski, and Y. Schüler. 2020. "On Adjusting the One-Sided Hodrick-Prescott Filter." Discussion Paper No. 11/2020, Deutsche Bundesbank.

Empirical Evidence on the Effectiveness of Capital Buffer Release*

Vasja Sivec[a] and Matjaž Volk[b]
[a]STATEC/ANEC
[a]Bank of Slovenia

With the new regulatory framework, known as Basel III, policymakers introduced a countercyclical capital buffer. Due to its recent introduction, empirical research on its effects is limited. We analyze a unique policy experiment to evaluate the effects of buffer release. In 2006, the Slovenian central bank introduced a temporary deduction item in the capital calculation, creating an average capital buffer of 0.8 percent of risk-weighted assets. It was released at the start of the financial crisis in 2008 and is akin to a release of a countercyclical capital buffer. We estimate its impact on bank behavior. After its release, firms borrowing from banks holding 1 pp higher capital buffer received 11 pp more in credit. Also, we find the impact was greater for healthy firms, and it increased loan loss provisioning for firms in default.

JEL Codes: G01, G21, G28.

*We gratefully thank Yi-An Chen for his invaluable contribution to this paper. We would also like to thank the participants of the ECBN conference Macroprudental Instruments and Financial Cycles (2016) and the anonymous referees that helped us to greatly improve the paper. Finally, we thank Karmen Kunčič from the Supervision department of the Bank of Slovenia for providing us detailed data on the prudential filter and for helping us understand its workings. The views expressed in this paper are solely the responsibility of the authors and do not necessarily reflect the views of the Bank of Slovenia or Statec and funding partners. Vasja Sivec gratefully acknowledges the support of the Observatoire de la Compétitivité, Ministère de l'Economie, DG Compétitivité, Luxembourg, and STATEC, the National Statistical Office of Luxembourg. Author e-mails: vasja.sivec@statec.etat.lu, matjaz.volk@bsi.si.

1. Introduction

In response to the financial crisis, regulators introduced several macroprudential instruments.[1] They are designed to impede the accumulation of systemic risk and to increase a bank's resilience to shocks. One of the key instruments introduced in Basel III is the countercyclical capital buffer (CCyB). In the periods of excessive credit growth and buildup of system-wide risk, banks are required to build a capital buffer (of up to 2.5 percent of risk-weighted assets) in the form of Common Equity Tier 1 capital. It is to be released in downturns to avoid regulatory capital requirements reducing credit growth, which could undermine the performance of the real economy and result in additional credit losses (Basel Committee on Banking Supervision 2015).

With the outbreak of COVID-19, banks supervised by the European Central Bank (ECB) were allowed to operate below the level of Pillar 2 Guidance (P2G) capital and capital conservation buffer (CCB) requirements.[2] These measures were further enhanced by the relaxation of the CCyB by national macroprudential authorities.[3] Such unprecedented relaxation of capital requirements intends to support lending and aims to mitigate second-round effects of the lockdown measures via the banking sector.

Unfortunately, these measures are recent and there is little evidence of their effectiveness. In the EU, CCyB was introduced in 2016 and COVID-19 marks its first release. Our paper provides empirical evidence on the effectiveness of capital buffer release following adverse economic conditions. We study a unique policy experiment that mirrors the workings of a capital buffer release at the start of the 2008 financial crisis in the Slovenian banking system. We study its impact on bank lending and loss-absorption capacity. Our findings

[1]For an overview of macroprudential policy and its tools, see Basel Committee on Banking Supervision (BCBS) (2010), Arnold et al. (2012), Galati and Moessner (2013), Claessens (2014), Cerutti, Claessens, and Laeven (2017), and Kahou and Lehar (2017).

[2]https://www.bankingsupervision.europa.eu/press/pr/date/2020/html/ssm.pr200312~43351ac3ac.en.html.

[3]Until July 2020, 13 EU countries at least partial released the CCyB and only 5 countries preserved a positive buffer (for more information see https://www.esrb.europa.eu/national_policy/ccb/html/index.en.html).

are favorable and support the actions undertaken by the policymakers and supervisors in response to the COVID-19 outbreak.

Current empirical research on CCyB relies on models that proxy the effects of CCyB by using changes in capital ratios.[4] This approach could be flawed. First, capital ratios are slow to adjust. CCyB release is sudden and generates a discontinuous shift in capital ratios. Second, changes in capital ratios are endogenous. They are subject to banks' own decisions. Endogenous capital changes may have a different effect on credit supply compared with an exogenous CCyB release. In contrast, we employ a policy experiment where the release of a capital buffer is exogenous concerning the Slovenian banking system.

In 2006 Slovenian banks adopted International Financial Reporting Standards (IFRS). Under the IFRS, the loan loss provisions were calculated differently than under the approach of the preceding Slovenian Reporting Standards. As a result, banks were allowed to hold fewer provisions. Being prudent, Bank of Slovenia (BS) required banks to use the difference in the amount of provisions as a deduction item in the calculation of the capital adequacy ratio. The deduction item was called the prudential filter. Due to it, banks held additional capital from 2006:Q1 to 2008:Q3. In response to the financial crisis, it was abdicated. It amounted to 0.8 percent of a system's risk-weighted assets (RWA) and acted like a countercyclical buffer. Banks accumulated capital in good times only to use it as a buffer for losses in bad times.

To investigate the effects of capital buffer release in distressed economies, we consider the Slovenian banking system. It was one of the most severely affected banking systems in Europe in the global financial crisis. By 2013 its share of non-performing loans (NPL) reached 25 percent for the corporate sector. According to Hartmann, Huang, and Schoenmaker (2018), this places it third according to recapitalization costs among European countries, making it suitable as a case study of buffer release in a distressed European banking system.

[4]Akram (2014) uses a vector error-correction model and Gross, Kok, and Żochowski (2016) a global vector autoregressive model. Noss and Toffano (2016) use sign restrictions to identify shocks in past data that match a set of assumed directional responses of other variables to future changes in aggregate bank capital requirements.

Our identification strategy follows Khwaja and Mian (2008). We estimate the difference in a firm's credit growth between two (or more) banks with different sizes of a prudential filter. Because we compare a firm's response across banks, firm-specific shocks such as demand or firm risk are absorbed by firm-fixed effects. Therefore, we control for loan demand, and the observed effect that we identify is unbiased and relates only to differences in the loan supply of banks with different capital buffers.

We found evidence that a higher capital buffer caused higher loan growth after the release. In our benchmark model, for the same firm borrowing from at least two different banks, credit growth was 5–11 percentage points (pp) higher in a bank with a 1 pp higher capital buffer before its release.[5] In addition, the probability of loan increase for a firm was 5.8 pp higher with a bank with 1 pp higher capital buffer. We also find that lending was directed towards less risky firms. Finally, we test if banks used additional loss-absorption capacity to increase provisions for defaulted borrowers. Coverage ratio increased by 8.6 pp more in banks with a 1 pp higher buffer, for firms that defaulted at the time of buffer release. We find strong evidence for stabilizing effects of capital buffers. Several robustness tests confirm the validity of our results.

Our findings complement theoretical and simulation-based models that argue in favor of capital buffers. Borsuk, Budnik, and Volk (2020) explore the role of capital buffers in containing the reduction of lending to the real economy during the COVID-19 crisis. Their analysis employs a large semi-structural model that connects banks and macroeconomy. They find that capital buffers lead to higher lending, with positive effects on gross domestic product (GDP) and lower credit losses. Aikman, Nelson, and Tanaka (2015) use a three-period model and Rubio and Carrasco-Gallego (2016) a dynamic stochastic general equilibrium (DSGE) model in which CCyB reduces excess credit buildup. Brzoza-Brzezina, Kolasa, and Makarski (2015) employ a DSGE model to show that CCyB mitigates credit

[5]Related literature investigates how lending is affected by capital increases (as opposed to its releases). An overview of empirical literature can be found in Dagher et al. (2016). In Dagher et al. (2016), a 1 pp higher capital decreases lending from 0.15 pp to 8 pp, depending on the model and horizon considered (see Tables 4A and 4B).

imbalances in the buildup phase; however, loan-to-value (LTV) restriction is shown to be more effective in this respect. We show that CCyB is effective in the release phase where LTV cannot be effective by definition. Tayler and Zilberman (2016) and Gersbach and Rochet (2017) employ a DSGE model to show that CCyB curbs credit cycles. Additional support is provided by Biu, Scheule, and Su (2017), who apply simulation techniques to show that a higher capital buffer reduces system-wide losses and therefore increases the resilience of the Australian banking system. Their simulation also shows that banks would limit credit supply in response to higher capital requirements. We in addition analyze how buffer affects lending and loan loss provisioning in the downturn phase.

Our paper is closest to Jiménez et al. (2017). Jiménez et al. (2017) offer valuable and rich insight from an instrument called dynamic provisions. They use exhaustive loan-level data to show that the release of dynamic provisions increased credit supply in Spain when the crisis hit. To our knowledge, Jiménez et al. (2017) and we are the only two research studies that use a policy experiment to estimate the effects of a CCyB release. An important difference is that the dynamic provisioning follows a formula, so banks can anticipate future releases better than in our experiment. In our experiment, the release is caused by a crisis that was unexpected and exogenous for Slovenian banks. In addition, we provide evidence on the interaction of loan loss provisioning and capital buffer, which is an unresearched mechanism of this instrument.

Our findings carry implications for policymakers and supervisors. We show that capital release increases bank lending in a crisis period. Further, we found that the increased lending was largely directed towards less risky firms, those without delays in loan repayments. This is helpful because it intensifies the positive effect of a capital buffer release on the real economy. An additional favorable effect is faster recognition of losses by banks. As shown by Beatty and Liao (2011) and van Wijnbergen and Homar (2014), fast recognition of losses make crises shorter and less intense. Our findings show that a capital buffer was effective at the beginning of the crisis as banks with higher reserve capital provisioned by more.

The paper is structured as follows. In the next section, we introduce the prudential filter and macroeconomic environment in Slovenia for the period in which it was active. Section 3 presents the methodological approach and data used for the analysis. Section 4 presents the results. Finally, Section 5 concludes the paper.

2. Prudential Filter

This section provides insights into the functioning of the prudential filter. The prudential filter was introduced at the beginning of 2006 and released at the end of 2008 when the crisis hit. We first discuss the macroeconomic and banking environment in Slovenia in the period 2007–10 and then the prudential filter.

2.1 Macroeconomic and Banking Environment

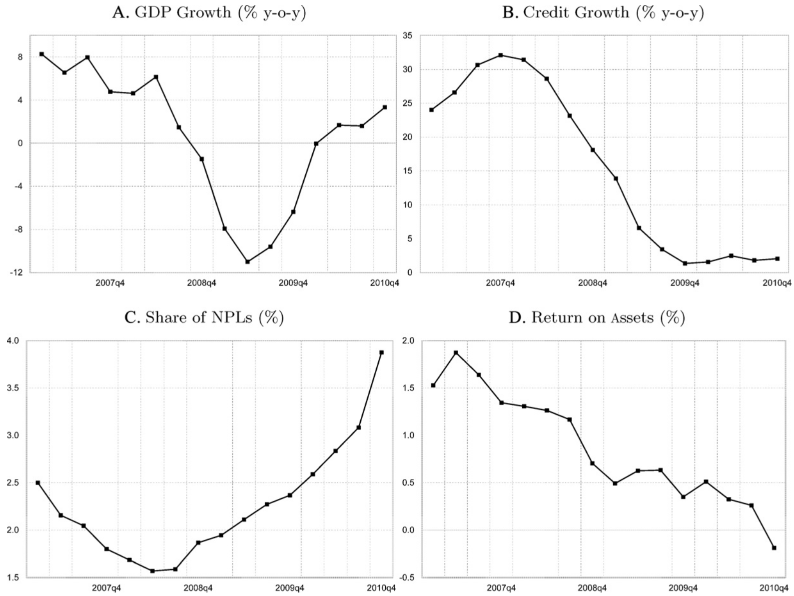
The period surrounding the buffer release is characterized by a transition from a period of high economic and credit growth to a deep recession. After a period of high growth, GDP turned negative in 2008:Q4 (see Figure 1).[6] At the time, the central bank of Slovenia released the prudential filter. In 2009 GDP contracted further, followed by a mild recovery in 2010. The recession severely affected the banking sector. Credit growth declined to 0 percent in 2009. A freeze of the European interbank market, which represented an important source of funding for Slovenian banks, contributed to this. A decrease in economic activity was accompanied by an increase in the share of non-performing loans. This latter became the main problem of Slovenian banks.[7] Concurrently, bank profit declined. In 2010 it turned negative and Slovenian banks started recording losses. Between 2009–14 these losses amounted to 10 percent of total pre-crisis assets.

The Bank of Slovenia decided to release the prudential filter in 2008:Q4. This was the time of the first signs of a banking crisis, triggered by an exogenous shock. A deep contraction of credit growth followed in 2009. It was accompanied by a decrease in economic activity that likely decreased loan demand. An estimation methodology that does not control for a fall in loan demand will lead to

[6]Banking-sector variables are calculated as weighted averages across banks. A bank's weight corresponds to a bank's share in total assets.

[7]In this study we define NPLs as loans to borrowers classified as C, D, or E in the five-grade rating scale from A to E.

Figure 1. Macroeconomic and Banking Environment in Slovenia in 2007–10



Source: Bank of Slovenia, own calculations.

a biased estimate because its decrease would attenuate the size of coefficients. Our identification strategy is free from this bias. We employ a loan-level differences-in-differences model to control for loan demand (see Section 3.1).

2.2 Functioning of Prudential Filter

Following the introduction of International Financial Reporting Standards (IFRS) in 2006, the Bank of Slovenia introduced the prudential filter. The prudential filter implicitly increased regulatory capital requirements, acting as CCyB. These requirements were released in 2008:Q4. This section describes the nature and regulatory aspects of the prudential filter.

In 2002 the European Parliament and Council adopted the Regulation EC/1606/2002. It required EU banks to traverse from national accounting standards to IFRS by January 2006. The Regulation had a major impact on the Slovenian banking sector. Under the IFRS

Table 1. Provision and Impairment Rates Valid in the Slovenian Banking Sector Before 2006

Rate	Credit Rating	Description
1%	A	Official institutions, no overdue, premium collateral
10%	B	Expected to be repaid, overdue under 30 days
25%	C	Insufficient cash flow, overdue 30–90 days
50%	D	Not expected to be repaid in full, overdue 90–360 days
100%	E	Not expected to be repaid, overdue above 360 days

Source: Provision or impairment rates can be found in the Official Gazette of the Republic of Slovenia (2005a, Article 22). Definitions of asset classes can be found in the same document, under Article 11.

accounting standards, provisions and impairments are recorded at fair value instead of at historical cost, as was done before 2006 under the Slovenian Accounting Standards.

A bank loan carries a risk that a borrower may not repay it. To account for such losses, banks apply impairments which are the difference between the carrying amount of the loan and the recoverable amount.[8] They are conventionally expressed in percentages of the carrying amount of the loan. A bank records the impaired value of the loan on the assets side of its balance sheet. On the liabilities side of the bank's balance sheet, impairments reduce the amount of capital. This is because the impaired amount of the loan enters into the bank's income statement as a deduction to the bank's profit, which is subsequently added to the bank's capital. The bottom line is that the higher/lower the impairments, the more/less capital a bank needs to hold to be compliant with regulatory capital requirements.

Before 2006, provisioning rates were set by the Bank of Slovenia. It set them based on historical data in a conservative manner. Provision and impairment rates applicable before 2006 are presented in Table 1. When the bank issued a loan, it immediately impaired the carrying amount in line with risk buckets presented in Table 1. If a

[8]This definition is derived from the official definition published in the Official Gazette of the Republic of Slovenia (2015).

loan was downgraded to a higher risk bucket, the bank had to apply a higher provision rate, irrespective of materialization of losses.

In 2006, Slovenian banks traversed to IFRS. Under the IFRS, provision and impairment rates were no longer set by the Bank of Slovenia. They were set by the banks using a fair-value approach. Many banks kept the system of assigning provisions based on credit ratings. But, importantly, banks were now free to determine provisioning rates for each risk bucket. They no longer applied those presented in Table 1.

On average, the historical approach imposed higher provision and impairment rates than the fair-value approach. Under the fair-value approach, a bank is required to provision for materialized losses. In contrast, under the historical approach, the loan loss provisions are recorded regardless of actual losses.

The Bank of Slovenia expected the amount of provisions and impairments to decrease under the IFRS (see Bank of Slovenia 2015). A substantial decrease of provisions and impairments would increase bank profit, which could be paid out in dividends, making banks less capitalized and riskier.

To mitigate the reduction in bank capital, the Bank of Slovenia amended the rules on credit risk calculation[9] and the regulation on bank capital calculation.[10] The amendments stated that, for regulatory purposes, the banks were required to introduce a (own funds) deduction item.[11] It was named *prudential filter* and was calculated as the difference between provisions and impairments calculated by using the historical approach rates and the provisions and impairments calculated under the fair-value approach. This rule applied only to loans and claims that were provisioned collectively under the IFRS. Individually impaired loans, which are to a large extent non-performing loans, were exempt from this calculation because for these loans a bank thoroughly assesses the expected cash flow and provisions accordingly.

Since prudential filter was deducted from Tier 1 capital, it forced banks to hold higher capital from 2006:Q1 to 2008:Q3. This

[9] See Official Gazette of the Republic of Slovenia (2005a).

[10] See Official Gazette of the Republic of Slovenia (2005b).

[11] Own funds is a broader definition of capital that also includes Tier I capital and secondary capital.

approximated the effect of a countercyclical buffer buildup if it existed at the time.

Our identification strategy relies on the filter release being caused exogenously by the financial crisis. To this purpose, the results include a placebo test. In addition, here we argue that the filter release remained uncertain until the start of the financial crisis in Slovenia.

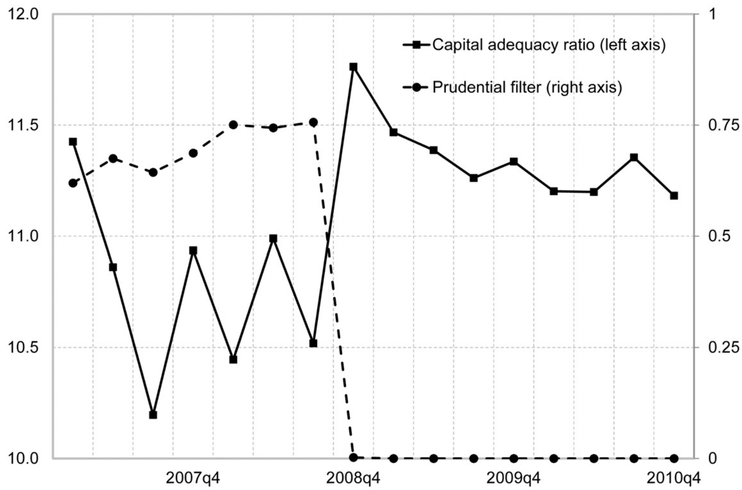
The special report of the Bank of Slovenia for the National Assembly on the causes of the capital shortfalls of banks[12] describes that "banks and auditors moved to have it ([prudential filter]) revoked several times" (Bank of Slovenia 2015). They were rebuffed until December 2007. Even then, the Bank of Slovenia deferred abdication "until a slowdown in the excessive lending activity of the banks [October 2008]" (Bank of Slovenia 2015). Uncertainty regarding its release is further supported by national newspaper articles. They report on the dissatisfaction of commercial banks with the deferment of the prudential filter abdication. In October 2008, a leading national newspaper reported that "at the beginning of the year [2008] the central bank governor . . . opposed abdication of the filter because of excess credit activity in the previous year (author's translation, Zimic 2008)." This points toward a substantial degree of uncertainty regarding the timing of the filter's abdication before its actual release in response to the financial crisis.

On several occasions, banks requested to abdicate the prudential filter. That would make banks more profitable per unit of capital, but also less resilient to future shocks. The Bank of Slovenia declined their requests and only abdicated the prudential filter in 2008:Q4, at the first signs of the financial crisis. As a direct impact of the abolishment of the prudential filter, the bank capital adequacy ratio increased, on average by 0.8 percentage point. Sudden increases in bank capitalization implied that banks could use excess capital for either lending or absorption of credit losses, which is analogous to a countercyclical capital buffer release.

The functioning of the prudential filter is presented in Figure 2. The dashed line shows the amount of the prudential filter, which was about 0.8 percent of RWA before the release and zero afterwards.

[12]In English, p. 41, section Introduction of an own funds deduction item.

Figure 2. Weighted Mean of Capital Adequacy Ratio and Prudential Filter in Percent of RWA, 2007–10



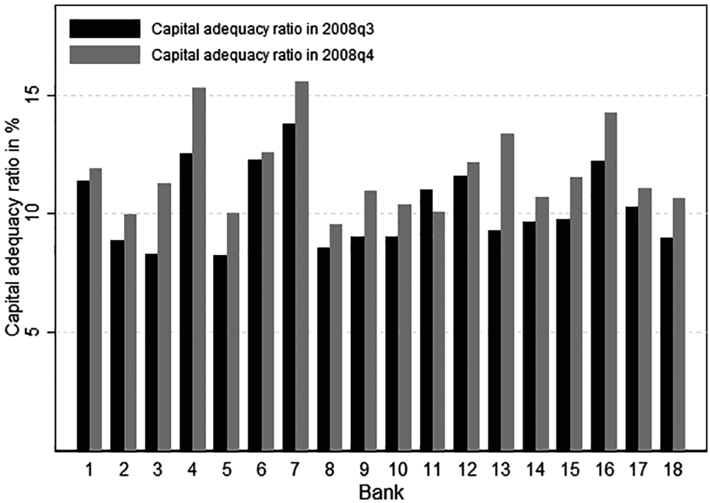
Source: Bank of Slovenia, own calculations.

The capital adequacy ratio (solid line in Figure 2) displays a mirrored picture. It increased almost one-to-one when the prudential filter was released. The prudential filter increased capital requirements during an expansionary period and alleviated them in the time of the financial crisis.

Figure 3 shows the capital adequacy ratio (CAR[13]) by banks before and after the release. The prudential filter caused an increase in the CAR for all banks except one. Note the difference between the dashed and solid line in Figure 3. It does not arise only due to a prudential filter release. There might have been other factors influencing the change in the CAR between 2008:Q3 and 2008:Q4, say recapitalization or realization of losses. This explains a decrease in the CAR for the one bank, which could not arise due to the prudential filter release. The prudential filter can only increase the capital available to a bank.

[13]Throughout this paper the acronym CAR stands for capital adequacy ratio and is not to be confused with cumulative abnormal returns.

Figure 3. Capital Adequacy Ratio Before the Release (2008:Q3) and After It (2008:Q4), Across Banks

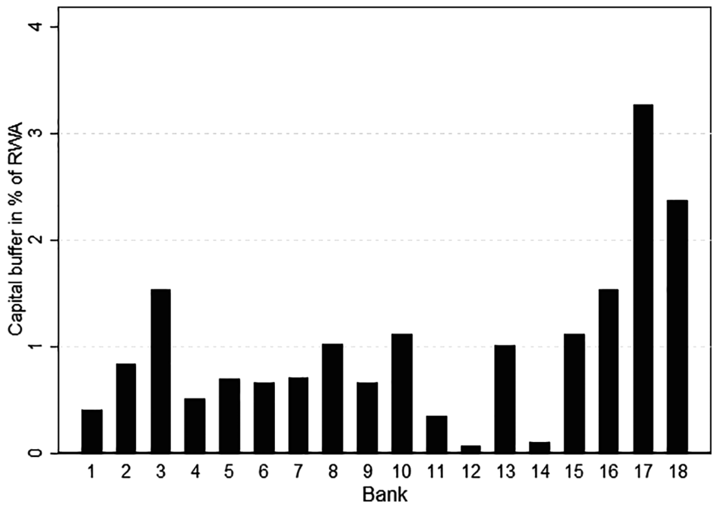


Source: Bank of Slovenia, own calculations.

Figure 4 shows the size of the prudential filter in terms of RWA before its release in 2008:Q3. We tested if banks that were required to hold a higher prudential filter lent and provisioned by more at the beginning of the crisis. Our identification strategy (described in Section 3.1) relies on firms taking loans with multiple banks subject to varied prudential filter requirements. Loan-level data coupled with between-bank prudential filter variability enabled us to estimate the effect of a 1 pp increase in the capital buffer on bank lending while controlling for loan demand. Eight banks held prudential filters above 1 percent and eight in the range of 0.3–1 percent of RWA. Two banks held prudential filters close to 0 percent. With release prudential filter translated into an increase in capital adequacy by the same amount.

There is a conceptual difference between the prudential filter and the CCyB. Under the CCyB the rate of additional capital is the same for all banks (up to 2.5 percent of RWA). On the other hand, the prudential filter was bank specific. It ranged from close to 0 percent of RWA to more than 3 percent. The fact that the prudential filter varied facilitates our analysis. Its variability across banks enabled

Figure 4. Prudential Filter in 2008:Q3 in Percent of RWA, Across Banks



Source: Bank of Slovenia, own calculations.

us to estimate the average effect of a 1 pp increase in the capital buffer. Note also that the CCyB is applied by increasing the minimum capital requirement, whereas the prudential filter decreased the accounting value of capital that entered the calculation of capital adequacy ratio. Regardless, in practice, they both increase the capital available to banks at the time of its release.

3. Methodology

We now present the identification strategy and data used to estimate the effect of the capital buffer release on bank lending and loan loss provisioning.

3.1 Identification Strategy

We identified the effects of buffer release in a loan-level model. Its key advantage is that it controls for loan demand and thereby yields unbiased and consistent estimates of coefficients. The methodology used in this section was put forward by Khwaja and Mian (2008). It was further adopted by Jiménez et al. (2010, 2017), Behn, Haselmann, and Wachtel (2016), Bonaccorsi di Patti and Sette (2016), and others.

Khwaja and Mian (2008) use a clever estimation technique that allows one to control for loan demand. Suppose that we have N borrowers with at least two banking relations in a given period:

$$y_{ij} = \beta X_{ij} + D_i + \epsilon_{ij}, \tag{1}$$

where y_{ij} stands for borrower i's loan growth ($i = 1 \dots N$) in bank j ($j = 1 \dots M$) in the period surrounding the buffer release (see Section 3.2). X_{ij} represents a $K \times 1$ vector of policy and control variables. D_i is a dummy variable that takes the value of 1 for individual i and 0 elsewhere. It absorbs firm-specific (unobservable) loan demand and other firm-specific characteristics. It enables us to estimate the effect of policy variable X_{ij} on loan growth y_{ij}, while controlling for unobservable firm-specific characteristics.

In our case, we estimate the effect of prudential filter release on a bank's loan supply and loan loss provisioning. In the first case the dependent variable is loan growth, and in the second it is the change in coverage ratio realized by bank j to firm i. Two key factors defining the rate of provisioning are firm riskiness and the amount of collateral. While both variables can in general be observed, our loan-level methodology is still advantageous. It captures all firm-level effects, including riskiness and availability of collateral. We address other potential firm-bank specific issues in Section 4.

3.2 Data

We used data from the credit register of the Bank of Slovenia. It contains multiple observations per individual borrower for each period. Having multiple observations per borrower allowed us to control for individual-specific fixed effects. Loans were obtained from the population of 18 banks. On the level of a borrower these are only available for firms. Households loans are reported cumulatively across risk buckets and cannot be used in a loan-level model. By considering only corporate loans, we still considered nearly all loans to the private non-financial sector. Loans to households represented only 23 percent of all credit to the private non-financial sector in 2008.[14]

The first important step in data preparation was to select an appropriate period to be used for calculating loan growth. Our baseline period is credit growth between one quarter before the prudential filter release (2008:Q3) and three quarters after the release (2009:Q3). One could argue that the chosen period is subjective. Therefore, we also estimated the model on horizons from one to four quarters after the release and report on those results.

For identification purposes, we restricted our sample to firms indebted to at least two banks. After imposing this restriction we were left with 7,882 firms. They account for 22.3 percent of all the firms that were in the same period indebted to at least one bank. Admittedly, this share is low. However, their total loan amount accounts for 84.2 percent of loans. Thus the data are representative and cover a large share of the total amount of lending tofirms. Next, for estimating the effect of buffer release on lending, we restricted our sample to performing firms alone. We excluded the non-performing firms because accounting rules dictate that non-paid interest on NPLs have to be added to the amount of non-performing loans. This increase in the loan amount is caused by accounting regulation and could be spuriously correlated with our regressors. Lastly, to eliminate outliers, we excluded firms of the 1st and 100th percentile of the distribution of our dependent variables.

In estimating the effect on loan loss provisioning, we focused on firms that are either in default or have difficulties in repaying the loan. Only these need to be provisioned extensively and account for the bulk of loan loss reserves. If we included the performing firms, we would find a much smaller or even insignificant effect on provisions. The reason is that there is no need to increase provisions for

[14]We also performed an aggregate analysis with loans to households included. We estimated a bank-level dynamic panel-data model with loan growth to firms and households as the dependent variable. The results are in line with the findings presented in Section 4. The estimated effect of buffer release on bank lending is, however, lower, which can be attributed to the lack of control for loan demand in the bank-level model, different sample, and different estimation methodology. The results are available upon request. We do not report on them because of the potential presence of omitted-variable bias.

Table 2. Summary Statistics

Variable	Unit	Mean	Std. Dev.	Min.	Max.	Obs.
Loan Growth	%	14.79	105.36	–90.47	1,166.67	11,984
Loan Increase	0/1	0.34	0.47	0	1	11,984
Change in Coverage Ratio	pp	11.30	28.46	–81.70	92.86	1,429
Prudential Filter	%	0.72	0.36	0.07	1.53	11,043
Total Assets	EUR bln.	5.16	5.33	0.02	15.10	11,984
Capital Adequacy Ratio	%	10.07	1.60	8.22	15.23	11,984
Share of NPLs	%	2.66	1.20	0.05	4.77	11,984
Bank Credit Growth	% y-o-y	25.72	13.73	9.80	54.08	11,984

Source: Bank of Slovenia, own calculations.
Note: Loan growth is calculated for the period 2008:Q3–2009:Q3. Credit increase is a dummy variable equal to 1 if firm i loan amount increased in bank j in period 2008:Q3–2009:Q3. Prudential filter, total assets, capital adequacy ratio, share of NPLs, and bank credit growth are reported at their values from 2008:Q3, just before the release took place. Change in coverage ratio is calculated for defaulted firms, whereas all other statistics are reported for performing part of the sample.

firms that repay loans regularly. This follows from the IFRS-incurred loss provisioning model. Similarly, as in the case of the loan growth analysis, we eliminated outliers.

Table 2 shows summary statistics for the variables included in the model. Credit growth is calculated as percentage growth in credit between one quarter before and the third quarter after its release (2008:Q3 to 2009:Q3). Mean credit growth is 15 percent. Loan increase is a dummy variable equal to 1 if firm i's loan amount increased in bank j in the period 2008:Q3–2009:Q3. Thirty-four percent of the firms increased their indebtedness after the release. The second variable of interest is the change in coverage ratio. It has a mean equal to 11.3 pp.[15] It is calculated only for the non-performing firms. All policy and control variables are included in the model at their values in 2008:Q3, i.e., just before the release. The average value of our main policy variable, the prudential filter, was

[15] Change in coverage ratio is calculated as $\Delta CR_{ij} = \frac{Provisions_{ij,2009q3}}{Loans_{ij,2009q3}} - \frac{Provisions_{ij,2008q3}}{Loans_{ij,2008q3}}$.

0.72 percent in 2008:Q3.[16] Bank size is measured with total assets. Its average value in 2008:Q3 was about EUR 5 billion. Average capital adequacy ratio, the share of non-performing loans, and year-over-year (y-o-y) bank credit growth before the filter release were 10.1 percent, 2.7 percent, and 25.7 percent, respectively.

4. Results

We now discuss the results. We investigated if banks with a higher amount of capital buffer lent more at the beginning of the crisis in 2009. Next, we explored the characteristics of firms that benefited from additional lending. Lastly, we verified if banks used extra loss-absorption capacity to increase provisioning for bad loans. By answering these questions, we evaluate the effectiveness of the capital buffer release policy.

Table 3 shows the effect of the buffer release on bank lending. The dependent variable is firm i credit growth for a loan taken with bank j in the period 2008:Q3–2009:Q3. We control for firm-specific demand with firm-fixed effects and include several controls for bank-level factors. Model 1 in Table 3 shows our baseline results. We find that for the same firm, borrowing from at least two different banks that differ in the size of the prudential filter, credit growth was 11.1 pp higher if the bank had a 1 pp higher capital buffer.[17] By using standard errors clustered at the bank level, this coefficient is statistically significant at conventional levels.[18] This implies that capital buffer release indeed increases bank lending.

We now extend our baseline model by adding the credit growth in the year before the prudential filter release. If banks that held a higher amount of prudential filter are the banks that lent more before the capital release, then the identified effect could be incorrectly attributed to the prudential filter. It might reflect higher credit growth of banks that incidentally also held higher prudential filters.

[16]This is a non-weighted average expressed from the restricted sample. Its maximum value is 1.53 percent, whereas this same value is more than 3 percent when expressed from the unrestricted sample. Figure 4 plots the unrestricted sample.

[17]Compared with the standard deviation of capital adequacy ratio, which is 0.36 pp, an average increase of 0.8 pp is considered substantial. If buffer instead increased by 0.36 pp, the loan growth would increase by 4 pp.

[18]Standard errors and p-values are corrected for small bank-level cluster size.

Table 3. The Effect of Capital Buffer Release on Bank Lending

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Prudential Filter	0.111**	0.118**	0.118**	0.124**	0.130**
	(0.049)	(0.051)	(0.044)	(0.048)	(0.057)
Capital Adequacy Ratio	0.016	0.020*	0.014	0.017	0.018
	(0.011)	(0.010)	(0.010)	(0.011)	(0.124)
Share of NPL	0.024*	0.032*	0.018	0.024	0.024
	(0.013)	(0.016)	(0.011)	(0.014)	(0.014)
Total Assets	–0.000*	–0.000	–0.000*	–0.000*	–0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Credit Growth		0.131			
		(0.167)			
Share of Interbank Funding			–0.122		
			(0.157)		
Prudential Filter*$I(Overdue > 0)$				–0.212**	
				(0.075)	
Prudential Filter*Rating					–0.048
					(0.032)
Constant	–0.124	0.231	–0.048	–0.134	–0.145
	(0.146)	(0.177)	(0.154)	(0.142)	(0.164)
Firm FE	Yes	Yes	Yes	Yes	Yes
No. of Observations	11,043	11,043	11,043	11,043	11,043

Source: Bank of Slovenia, own calculations.
Note: The table reports the estimation results for the loan-level differences-in-differences model. The dependent variable in all the equations is firm i loan growth in bank j in period 2008:Q3–2009:Q3 (10 percent is expressed as 0.1). Prudential filter is its amount in 2008:Q3 (just before the release), expressed in percent of RWA. Capital adequacy ratio, share of NPL, bank total assets, and share of interbank funding are taken from 2008:Q3. Credit growth is bank-specific credit growth in the year before the prudential filter release. $I(Overdue > 0)$ is an indicator equal to one if firm i repays the loan to bank j with overdue higher than zero days. Rating is a credit rating assigned by bank j to firm i and takes values from 0 (rating A) to 4 (rating E). Standard errors (in parentheses) are clustered at the bank level. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

As shown with model 2 in Table 3, the results are robust. Even when controlling for a bank's past credit growth, the prudential filter displays a positive and statistically significant effect. In addition, the effect of bank credit growth before the release of capital is found to be insignificant.

Next specification controls for the simultaneity of interbank credit market freeze and the buffer release. Before its release, the

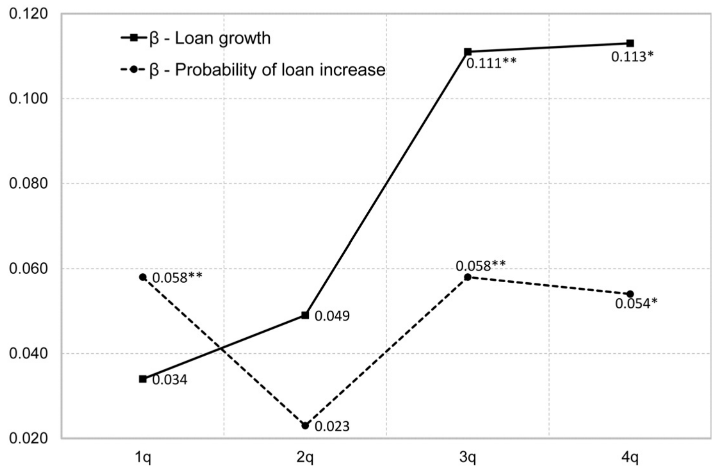
Slovenian banking system relied on interbank funding. Interbank funding increased from 10 percent in 2002 to approximately 40 percent of total funding in 2008:Q3. This share increased to 55 percent for foreign-owned banks. With the Lehman Brothers bankruptcy, the interbank market froze. This exogenous supply shock coincides with the timing of the buffer release. Failing to control for it could induce a bias in our estimates. Hence, specification 3 controls for the share of interbank financing. Despite controlling for it, the coefficient on the prudential filter remains the same in magnitude and statistical significance.

Our next set of results investigates which firms benefited from the positive effect of the filter release. Note that this was a period when the crisis began and non-performing loans started to accumulate. In response to it, banks could engage in evergreening of riskier loans. This practice reduces the pressure of loan loss provisions on bank capital and is documented in Peek and Rosengren (2005). It would be undesirable for the capital buffer release to amplify this effect.

We verified this by interacting prudential filter with two variables that measure firm riskiness. First, we used the number of days overdue in loan repayment by firm i to bank j. Model 4 in Table 3 shows this result. The interaction term is negative. In addition, the sum of the coefficients for a prudential filter and the interaction term is also negative. This implies that the positive effect of the prudential filter release is not only reduced for borrowers that have difficulties with loan repayment but is even negative. Second, we used the credit rating assigned by bank j to firm i. It takes a value from 0 (rating A) to 4 (rating E). The coefficient on the interaction term in specification 5 is negative, although it is not statistically significant (exact p-value is equal to 0.153). Overall, we conclude that solid and safe firms gain the most from a capital buffer release. This is an outcome desired by the policymakers.

We have shown that the capital buffer release increased loan growth in a specific time horizon, 2008:Q3–2009:Q3. We now verify the robustness of the results to the chosen time horizon. 2008:Q3 was used as a cut-off date before the prudential filter release. We kept this date fixed to stay as close as possible to the time of the buffer release. We did not wish to contaminate the dependent variable with other effects that transpired. For the same reason, we did not consider periods beyond one year after its release.

Figure 5. Coefficient for Loan Growth and for the Probability of a Loan Increase on One- to Four-Quarter Horizon After the Release



Source: Bank of Slovenia, own calculations.
Note: Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 5 presents the results for horizons that span from one to four quarters after the release. The effect of capital release on loan growth peaked in the third quarter after the release. Importantly, the estimated coefficient is positive in all the cases. It is, however, statistically significant only for the third and fourth quarter after the release.[19] This is to some extent expected since banks typically need time to reallocate spare capital.

We also estimated the probability of a loan increase following the release of the prudential filter. The dependent variable is equal to 1 if firm i's amount of loan borrowed from bank j has increased in the examined period. We used the same time horizon as in our benchmark regression. The advantage of this approach is that the estimated effects are not driven by outliers, which might be, despite certain exclusions, still present. The results are presented in Figure 5. Comparable to credit growth, the release of the capital buffer increases the probability of loan growth. We find that a firm

[19] Beyond the fourth quarter, the effects of buffer release is diminishing.

Table 4. The Effect of Capital Buffer Release on Bank Loan Loss Provisioning

	(1)	(2)	(3)
	Overdue$_{2009q3}$ > 0	**Overdue$_{2009q3}$ > 90**	**Overdue$_{2008q3}$ = 0, Overdue$_{2009q3}$ > 0**
Prudential Filter	0.084**	0.077**	0.086*
	(0.031)	(0.029)	(0.047)
Capital Adequacy Ratio	0.001	0.012	0.021
	(0.010)	(0.008)	(0.021)
Share of NPL	–0.014	–0.007	–0.014
	(0.016)	(0.009)	(0.022)
Total Assets	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Overdue	0.000	0.000	0.001***
	(0.000)	(0.000)	(0.000)
Constant	–0.026	0.006	–0.235
	(0.098)	(0.092)	(0.233)
Firm FE	Yes	Yes	Yes
No. of Observations	2,032	1,337	856

Source: Bank of Slovenia, own calculations.
Note: The table reports the estimates for the loan-level differences-in-differences model. The dependent variable in all the equations is the change in loan loss provisioning ratio between 2008:Q3 and 2009:Q3. The model is estimated for three subsamples: (1) and (2) include firms that had an overdue higher than 0 or 90 days, respectively, whereas (3) includes firms that were in overdue for the first time after the buffer's release. Prudential filter is recorded at its amount in 2008:Q3 (just before the release) and expressed in percent of RWA. Capital adequacy ratio, share of NPL, and bank total assets are taken from 2008:Q3. Overdue controls for firm i's overdue in bank j. Standard errors (in parentheses) are clustered at the bank level. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

had a 5.8 pp higher probability of a loan increase with a bank that held a 1 pp higher capital buffer.

We next explored the effect of capital release on bank loan loss provisioning. Due to a filter release, banks obtained spare capital. Spare capital increased their loss absorption capacity. However, a study by Brezigar-Masten, Masten, and Volk (2015) showed that banks intentionally underestimated the loan loss provisions when the non-performing loans started piling up in their balance sheets. We tested if banks with higher capital buffers provisioned more, thereby decreasing the underestimation of credit risk.

Results are presented in Table 4. The dependent variable is the change in the coverage ratio for each observation between 2008:Q3 and 2009:Q3. We controlled for firm fixed effects and focused on firms that were either in default or overdue. We present three different sets of results that depend on the firm's overdue.

We found that the prudential filter release increased loan loss provisioning. Model 1 in Table 4 shows the results for the sample of firms that were past due with loan repayment in 2009:Q3 for at least one day. For the same firm, the coverage ratio increased by 8.4 pp more with banks that held a 1 pp higher capital buffer. Next, we used stricter criteria in sample selection. We included only firms that were more than 90 days overdue. This threshold is typically used to classify borrowers as non-performing, so banks provision extensively only after it is bridged. The results, presented in column 2, confirm our previous findings. One might be concerned that for the majority of firms included in models 1 and 2 the coverage ratio is constant. This could be because these firms were in default for a period that was long enough to be fully provisioned for. The average number of days overdue among defaulted firms is above 500 days. To address this issue, we estimated a model for firms that became past due after the prudential filter abdication. These were new defaulters that banks provisioned for the first time after capital release. The results are in column 3 of Table 4. As before, we find a stable and positive effect similar in magnitude to our previous results. We reconfirm that the buffer release increased the loss absorption of banks, as intended by the policymakers.

We now address some firm-bank characteristics that could influence our finding of increased provisioning following a buffer release. The longer the time in default, the higher should be the coverage ratio of a loan. Firms, however, do not start to delay loan repayment to all banks at the same time. There might be a difference in the coverage ratio for the same firm across multiple banks. To address this we add overdue-in-loan-repayment as a control variable. For models 1 and 2 this control is irrelevant. The difference in overdue of 10 or 50 days is negligible for firms that have been overdue for a long time. Once the number of days in overdue becomes high, banks estimate that it is unlikely that a loan will be repaid and they provision accordingly. For new defaulters, as in model 3, this variable is found to be relevant. A firm that started to delay loan repayment with bank A 50 days before it started to delay loan repayment with bank B is expected to have on average a 5 pp higher coverage ratio in bank A as compared with bank B.

The second determinant of loan loss provisioning is collateral. Omission of collateral is to some degree controlled for by the fixed effects. They capture the firm's total collateral. Banks, however, differ in strategy and ability to engage a firm's collateral. Unfortunately, we cannot control the exact amount of collateral pledged by firm i in bank j. These data are not available. We instead assess the direction of bias assuming that collateral does affect loan loss provisioning.

The direction of (potential) bias, due to omission of collateral, will depend on the correlations between provisioning, collateral, and the prudential filter. First, we establish that the prudential filter and collateral are positively correlated. Banks that held higher filters incurred lower losses in 2009–14.[20] Next, we know that collateral and loan loss provisions are also negatively correlated. This follows basic accounting rules. Had loans been fully collateralized, there would be no need for provisions. Finally, if prudential filter acts as a proxy for collateral, the coefficient is expected to be downward biased. Our estimates of the effect of the capital buffer on provisioning represent a lower boundary on the coefficient estimate.

4.1 Robustness Checks

This section presents four sets of robustness checks.[21] First, we expand the sample by adding single-bank relation firms to (potentially) increase the external validity of our results. Second, we conduct a placebo test. It rules out that our results are driven by a particular set of confounding factors. Third, we evaluate the robustness of our results to unobserved confounders by using the bias adjustment approach of Oster (2019). And fourth, we apply a matching estimator to control for potential data imbalance and to control for variables that were omitted from the response regression (loan growth) but affect treatment assignment (buffer size). Robustness checks are applied to our benchmark model 1 in Table 3.

[20] The correlation between bank losses in 2009–14, expressed in terms of precrisis assets, and prudential filter in 2008:Q3 (in percent of RWA) is equal to –0.3.

[21] We thank the anonymous referees for their suggestions.

4.1.1 Sample Expansion

Khwaja and Mian (2008) control for firm-specific loan demand by relying on firms that borrow from multiple banks.[22] Single-bank firms are omitted.[23] Degryse et al. (2019) introduce an approach that does not rely on the presence of multiple banking relations. They replace the firm-specific dummy with a set of location-industry-size dummies.[24] Their model is identified if there are at least two firms in each location-industry-size bucket, regardless of the number of banking relations per firm.

The model of Degryse et al. (2019) requires stronger assumptions on firm-specific loan demand. Khwaja and Mian (2008) require loan demand of a single firm to be constant across banks, whereas Degryse et al. (2019) require it to be constant across all firms in the same location-industry-size bucket and the banks that they borrow from. Which method delivers externally valid estimates depends on the trade-off between the importance of the inclusion of single-bank firms versus the plausibility of the constancy of firm loan demand within location-industry-size buckets.

Table 5 presents the results. Following Degryse et al. (2019) we use two-digit NACE codes to form industry clusters (83), municipalities to form location clusters (213), and the number of persons employed to form size clusters (16).[25] This produces a total of 13.108 clusters which are granular enough to capture firm-specific loan demand.[26] We find that a 1 pp higher capital buffer increases loan growth by 9 pp. This is consistent with, but smaller than, our benchmark estimate (11.1 pp). We cannot say for certain if the marginally smaller estimate is the result of sample expansion or due to a less precise control for loan demand. However, both estimates of credit growth are substantial in economic terms.

---

[22]See section on identification strategy.

[23]These firms account for 79 percent of all non-financial corporations in our sample.

[24]In a dynamic setting their dummy variable is also time specific.

[25]These are internal Bank of Slovenia size clusters which are based on the Eurostat size clusters but are more detailed.

[26]We also estimated models with location clusters defined by regions (13), less granular size clusters (4), and industry clusters defined by one-digit NACE codes (20). The estimate of the coefficient on capital buffer ranged between 0.075 and 0.090 and always remained significant.

Table 5. The Effect of Capital Buffer Release on Bank Lending with Sample Expansion

	(1)
Prudential Filter	0.090**
	(0.039)
Capital Adequacy Ratio	–0.011
	(0.010)
Share of NPL	0.025
	(0.015)
Total Assets	–0.000
	(0.000)
Constant	0.126
	(0.119)
ILS FE	Yes
No. of Observations	36,708

Source: Bank of Slovenia, own calculations.
Note: The table reports the estimation results for the loan-level model of Degryse et al. (2019). Dependent variable is firm i loan growth in bank j in period 2008:Q3–2009:Q3 (10 percent is expressed as 0.1). Prudential filter is its amount in 2008:Q3 (just before the release), expressed in percent of RWA. Capital adequacy ratio, share of NPL, bank total assets, and share of interbank funding are taken from 2008:Q3. Credit growth is bank-specific credit growth in the year before prudential filter release. ILS FE stands for industry-location-size fixed effects. Standard errors (in parentheses) are clustered at the bank level. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4.1.2 Placebo Test

In a placebo test we falsely assume that the buffer was released a year before its actual release, in 2007:Q4. In that period Slovenian economy recorded record GDP growth and was not affected by the crisis. This test verifies if the treatment effects (buffer release) were present before the policy change took place. Should the effects be present before the policy change took effect, it would indicate that our results are driven by confounding factors correlated with subsequent capital buffer release. It could also signal that buffer release was anticipated in advance.

Due to data limitations, we depart from the benchmark model (see specification 1 in Table 3). Capital adequacy is only available

Table 6. Placebo Test—Hypothetical Buffer Release in 2007:Q4

	Actual Release in 2008:Q4	**Hypothetical Release in 2007:Q4**
Prudential Filter	0.093*	0.007
	(0.052)	(0.095)
Leverage Ratio	0.882	0.358
	(0.626)	(1.268)
Share of NPL	0.025	–0.000
	(0.145)	(0.025)
Total Assets	–0.000	–0.000
	(0.000)	(0.000)
Constant	–0.041	0.432***
	(0.124)	(0.126)
Firm FE	Yes	Yes
No. of Observations	11,043	10,141

Source: Bank of Slovenia, own calculations.
Note: The table compares the estimation results of the actual and the placebo experiment. The dependent variable is loan growth in the one-year window around the treatment date (2008:Q3–2009:Q3 for real experiment and 2007:Q3–2008:Q3 for placebo test). All the control variables are dated one quarter before the treatment date. Standard errors (in parentheses) are clustered at the bank level. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

from 2008 onward, hence we replace it with leverage ratio expressed as the book value of capital to total assets.

Table 6 displays the results. To ensure that variable substitution does not affect the test, we first replicate our benchmark regression with the capital adequacy ratio replaced by the leverage ratio and keep the buffer release in 2008:Q4. The coefficient on the prudential filter remained of similar magnitude and statistical significance.

We then assumed a counterfactual buffer release in 2007:Q4. The second column displays the results. The coefficient on the prudential buffer is close to zero and statistically insignificant (the exact p-value is 0.945). We conclude that it is unlikely that a correlated (omitted) confounder or the anticipation of buffer release drives our results.

Table 7. Robustness to Omitted-Variable Bias

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline Model $\widetilde{\boldsymbol{\beta}}$	**Zero Model $\mathring{\boldsymbol{\beta}}$**	**$\boldsymbol{\beta^*}$ with $\boldsymbol{\delta = -1}$**	**$\boldsymbol{\beta^*}$ with $\boldsymbol{\delta = 1}$**	**Within Conf. Interval?**	**$\boldsymbol{\delta}$ for $\boldsymbol{\beta^* = 0}$**
β R^2	0.111 0.592	0.061 0.000	0.075	0.146	Yes	–3.158

Source: Bank of Slovenia, own calculations.
Note: The table shows the bounds of the estimated effect of the prudential filter on lending using Oster (2019) methodology. The restricted model includes the same set of controls as specification (1) in Table 3. Zero model includes only prudential filter and intercept. The bounds are calculated assuming $R_{\max} = 1$.

4.1.3 Bias Adjustment for Selection on Unobservables

Based on Altonji, Elder, and Taber (2005), Oster (2019) developed a test for assessing bias from unobservable factors. The idea is to compare the coefficient on prudential filter from a regression with the full set of controls ($\widetilde{\beta}$) with the coefficient from a regression with intercept only ($\mathring{\beta}$). This delivers a *bounded estimator* ($\beta^\star$) defined as follows:

$$\beta^\star \approx \widetilde{\beta} - \delta(\mathring{\beta} - \widetilde{\beta})\frac{R^2_{max} - \widetilde{R}^2}{\widetilde{R}^2 - \mathring{R}^2}. \tag{2}$$

The difference in the two coefficients ($\mathring{\beta} - \widetilde{\beta}$) is rescaled by the difference in R^2 of the two regressions and expressed in relation to the highest possible value of R^2 (R^2_{max}). The latter cannot be identified and is replaced by R^2_{max} as $min(2.2 \times \widetilde{\beta}, 1)$ (see Oster 2019), which is $R^2_{max} = 1$ in our model. δ determines the degree of proportionality between selection on observables and unobservables and is set to $\delta = 1$ or $\delta = -1$, depending on the direction of bias. For practical reasons, we focus on $\delta = -1$, as it implies an upward bias in the coefficient on the prudential filter. We also express the needed strength of omitted factors, relative to control variables, that would reduce the value of the coefficient to zero. An implausibly high value implies a low likelihood of omitted factors.

Table 7 shows the results. Columns 1 and 2 report baseline and intercept only (denoted as “Zero Model”) inputs for the calculation of the bounded estimator. Column 3 reports the lower bound of the coefficient when $\delta = -1$. The value of the coefficient reduces compared with the baseline model. However, it remains relatively high and significant. It is unlikely that relevant unobservable variables were omitted from our benchmark regression. In fact, column 6 shows that the effect of omitted variables would have to be more than three times larger to reduce the coefficient of prudential filter to zero. Finally, bounded estimators (with $\delta = -1$ and $\delta = 1$) are enclosed within the confidence interval for $\widetilde{\beta}$. We conclude that it is unlikely that an omitted-variable bias affects our benchmark regression to a significant degree.

4.1.4 Propensity Score Regression

Propensity score models were developed to control for selection biases in non-experimental settings. They alleviate bias when a variable, important for selection into treatment (buffer size), is omitted from response regression.[27] They can reduce the imbalance in covariates in case of a "lack of complete overlap" and can produce a better estimate of the average treatment effect when the response to treatment is heterogeneous.

Generalization of the propensity score methods for the case of continuous treatment was introduced by Hirano and Imbens (2004). Further discussion and implementation can be found in Bia and Mattei (2008). Similar settings have been investigated by Arpino and Mealli (2011), Schuler, Chu, and Coffman (2016), Kim, Paik, and Kim (2017), or Zhou et al. (2020). None of them considers a setting such as ours. It simultaneously includes continuous treatment, hierarchically structured data (firm- and bank-level clusters), treatment assignment at a different cluster level (bank level) than the response regression (loan level) and a differences-in-differences model. We employed a Monte Carlo verification of its small sample properties to assess its suitability for a regression setting such as ours.[28]

[27] Due to perhaps being insignificant in sample data. In propensity score literature this property is called "doubly robust." It states that the propensity score model is unbiased if at least one of the two models, outcome regression or treatment assignment regression, is specified correctly.

[28] For space considerations the results and discussion are available upon request.

The propensity score model is estimated in two steps. We first estimate the probability of being "selected" into treatment (probability of being assigned a certain buffer size). We employed a top-down strategy to select regressors.[29] This model produces probability weights for "selection into treatment." Their inverses are used as regression weights in the second step when we estimate the loan growth model. Weighting creates a synthetic sample in which over-represented data are down-weighted and under-represented units are up-weighted. This delivers a sample in which confounders (included in the first-step regression) are orthogonal to the treatment. It renders regression unbiased due to those covariates included in the first step of the model.

Table 8 repeats the benchmark regression (1) and displays the results from two propensity score models 2–3. In propensity score model 2, selection into treatment is derived from the capital adequacy ratio, share of interbank loans, share of credit in the bank balance sheet, and the share of impairments. Credit in the bank balance sheet and share of impairments are not included in our benchmark model,[30] making the results reported here robust to these two confounders. We notice that the regression coefficient on capital buffer decreased to 5.4 percent but remains significant at the 10 percent level (p-value is 7.3 percent). We consider this to be the lower bound on the effect of capital buffer release on loan growth.

We also noticed that regression (2) creates an imbalance in capital adequacy ratio and credit share.[31] We now reestimate the propensity score model but only balance it for these two variables. Column 3

---

[29]We regressed the buffer on all regressors in the benchmark model and added additional regressors that could affect it. We then removed the least significant regressor one at a time and stopped when the remaining regressors were significant at the 10 percent level. Results remain unchanged at the 5 percent significance level. Regression results, balance plots, and correlation coefficients are available upon request.

[30]Due to being insignificant.

[31]Weighting the data with propensity score weights intends to "orthogonalize" pertinent covariates relative to the treatment variable (buffer). If successful, the correlation coefficient between buffer and covariates should approach zero. In practice, it is difficult to orthogonalize the sample for four variables simultaneously. When applying propensity score weights, capital adequacy ratio and credit share remained correlated with the buffer.

Table 8. Propensity Score Model

	(1)	**(2)**	**(3)**
Prudential Filter	0.111**	0.054*	0.084**
	(0.049)	(0.034)	(0.034)
Capital Adequacy Ratio	0.016	0.027**	0.038**
	(0.011)	(0.008)	(0.008)
Share of NPL	0.024*	0.003	0.008
	(0.013)	(0.008)	(0.008)
Total Assets	–0.000*	–0.000**	–0.000***
	(0.000)	(0.000)	(0.000)
Constant	–0.124	0.011	0.033***
	(0.146)	(0.007)	(0.007)
Firm FE	Yes	Yes	Yes
No. of Observations	11,043	11,043	11,043

Source: Bank of Slovenia, own calculations.
Note: The table reports the estimation results for the loan-level differences-in-differences model (1) and two propensity score models, (2) and (3). The dependent variable in all the equations is firm i loan growth in bank j in period 2008:Q3–2009:Q3 (10 percent is expressed as 0.1). Prudential filter is recorded at its amount in 2008:Q3 (just before the release) and expressed in percent of RWA. Capital adequacy ratio, share of NPL, and bank total assets are taken from 2008:Q3. Standard errors (in parentheses) are clustered at the bank level. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

in Table 8 displays the results. The coefficient on buffer increased to 8.4 percent and was significant at the 1 percent level.

Regardless of the variables used to model selection into treatment, the effect of buffer release on loan growth is positive and statistically as well as economically significant.

5. Conclusion

This paper studies a unique experiment in the Slovenian banking system in 2007–10. The experiment is called the prudential filter and it acted like a countercyclical capital buffer. In 2008:Q4, an exogenous shock caused the prudential filter abdication. This resulted in a one-time increase of bank capital by 0.8 percent of risk-weighted assets on average. We estimate how this release of bank capital, akin to a countercyclical capital buffer, affected the banking system at the start of the financial crisis.

Our key results are the following. First, we show that banks with larger capital buffers lend more. A firm borrowing from a bank with a 1 pp higher capital buffer recorded a 5–11 pp higher credit growth, depending on the preferred model. This result is robust to various model specifications, estimation horizons, and robustness checks for omitted-variable bias. Second, healthy firms benefited the most from the excess credit capacity created by the buffer release. This intensifies the positive effect of the buffer on the real economy. Finally, we show that banks used extra loss-absorption capacity, resulting from the buffer release, to provision more for defaulted borrowers. Since a delay in loan loss recognition prolongs and intensifies the effects of financial crises, the CCyB can be considered as an effective mitigation policy.

Our findings are important for policymakers, supervisors, and regulators. We show that capital-based macroprudential measures, such as capital buffer, are an effective tool to support lending in turbulent economic conditions. In addition, they increase the loss-absorption capacity of banks. Banks use it to provision more for non-performing exposures. In light of recent policy measures taken in response to the outbreak of the COVID-19, we expect capital measures to perform well. It will be interesting to contrast our results with ex post verification of capital measures that are now being taken in response to the COVID-19 outbreak.

References

Aikman, D., B. Nelson, and M. Tanaka. 2015. "Reputation, Risk-Taking, and Macroprudential Policy." *Journal of Banking and Finance* 50 (January): 428–39.

Akram, Q. F. 2014. "Macro Effects of Capital Requirements and Macroprudential Policy." *Economic Modelling* 42 (October): 77–93.

Altonji, J. G., T. E. Elder, and C. R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1): 151–84.

Arnold, B., C. Borio, L. Ellis, and F. Moshirian. 2012. "Systemic Risk, Macroprudential Policy Frameworks, Monitoring Financial Systems and the Evolution of Capital Adequacy." *Journal of Banking and Finance* 36 (12): 3125–32.

Arpino, B., and F. Mealli. 2011. "The Specification of the Propensity Score in Multilevel Observational Studies." *Computational Statistics and Data Analysis* 55 (4): 1770–80.

Bank of Slovenia. 2015. "Report of the Bank of Slovenia on the Causes of the Capital Shortfalls of Banks and the Role of the Bank of Slovenia as the Banking Regulator in Relation Thereto, the Recovery of Banks in 2013 and 2014, the Efficiency of the System of Corporate Governance of Banks under State Ownership and the Manner of Resolving the Consequences of the Capital Inadequacy of Commercial Banks." Technical Report. https://www.bsi.si/en/publications/recovery-of-banks-201314/report-of-the-bank-of-slovenia-for-the-national-assembly-march-2015.

Basel Committee on Banking Supervision. 2010. "Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems." Technical Report, Bank for International Settlements.

———. 2015. "Frequently Asked Questions on the Basel III Countercyclical Capital Buffer." Technical Report, Bank for International Settlements.

Beatty, A., and S. Liao. 2011. "Do Delays in Expected Loss Recognition Affect Banks' Willingness to Lend?" *Journal of Accounting and Economics* 52 (1): 1–20.

Behn, M., R. Haselmann, and P. Wachtel. 2016. "Procyclical Capital Regulation and Lending." *Journal of Finance* 71 (2): 919–56.

Bia, M., and A. Mattei. 2008. "A Stata Package for the Estimation of the Dose-Response Function through Adjustment for the Generalized Propensity Score." *Stata Journal* 8 (3): 354–73.

Biu, C., H. Scheule, and E. Wu. 2017. "The Value of Bank Capital Buffers in Maintaining Financial System Resilience." *Journal of Financial Stability* 33 (December): 23–40.

Bonaccorsi di Patti, E., and E. Sette. 2016. "Did the Securitization Market Freeze Affect Bank Lending during the Financial Crisis? Evidence from a Credit Register." *Journal of Financial Intermediation* 25 (January): 54–76.

Borsuk, M., K. Budnik, and M. Volk. 2020. "Buffer Use and Lending Impact." *Macroprudential Bulletin* (European Central Bank) (October): Article 11.

Brezigar-Masten, A., I. Masten, and M. Volk. 2015. "Discretionary Credit Rating and Bank Stability in a Financial Crisis." *Eastern European Economics* 53 (5): 377–402.

Brzoza-Brzezina, M., M. Kolasa, and K. Makarski. 2015. "Macroprudential Policy and Imbalances in the Euro Area." *Journal of International Money and Finance* 51 (February): 137–54.

Cerutti, E., S. Claessens, and L. Laeven. 2017. "The Use and Effectiveness of Macroprudential Policies: New Evidence." *Journal of Financial Stability* 28 (February): 203–24.

Claessens, S. 2014. "An Overview of Macroprudential Policy Tools." IMF Working Paper No. 14/214.

Dagher, J., G. Dell'Ariccia, L. Laeven, L. Ratnovski, and H. Tong. 2016. "Benefits and Costs of Bank Capital." Staff Discussion Note No. 16/04, International Monetary Fund.

Degryse, H., O. De Jonghe, S. Jakovljević, K. Mulier, and G. Schepens. 2019. "Identifying Credit Supply Shocks with Bank-Firm Data: Methods and Applications." *Journal of Financial Intermediation* 40 (October): Article 100813.

Galati, G., and R. Moessner. 2013. "Macroprudential Policy: A Literature Review." *Journal of Economic Surveys* 27 (5): 846–78.

Gersbach, H., and J.-C. Rochet. 2017. "Capital Regulation and Credit Fluctuations." *Journal of Monetary Economics* 90 (October): 113–24.

Gross, M., C. Kok, and D. Żochowski. 2016. "The Impact of Bank Capital on Economic Activity—Evidence from a Mixed-Cross-Section GVAR Model." ECB Working Paper No. 1888.

Hartmann, P., H. Huang, and D. Schoenmaker. 2018. "The Changing Fortunes of Central Banking." Bruegel Special Report.

Hirano, K., and G. W. Imbens. 2004. "The Propensity Score with Continuous Treatments." In *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives,* ed. A. Gelman and X.-L. Meng, 73–84. John Wiley & Sons.

Jiménez, G., A. R. Mian, J.-L. Peydró, and J. Saurina. 2010. "Local versus Aggregate Lending Channels: The Effects of Securitization on Corporate Credit Supply in Spain." Technical Report, National Bureau of Economic Research.

Jiménez, G., S. Ongena, J.-L. Peydró, and J. Saurina. 2017. "Macroprudential Policy, Countercyclical Bank Capital Buffers, and Credit Supply: Evidence from the Spanish Dynamic Provisioning Experiments." *Journal of Political Economy* 125 (6): 2126–77.

Kahou, M. E., and A. Lehar. 2017. "Macroprudential Policy: A Review." *Journal of Financial Stability* 29 (April): 92–105.

Khwaja, A. I., and A. Mian. 2008. "Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market." *American Economic Review* 98 (4): 1413–42.

Kim, G.-S., M. Cho Paik, and H. Kim. 2017. "Causal Inference with Observational Data under Cluster-Specific Non-ignorable Assignment Mechanism." *Computational Statistics and Data Analysis* 113 (September): 88–99.

Noss, J., and P. Toffano. 2016. "Estimating the Impact of Changes in Aggregate Bank Capital Requirements on Lending and Growth During an Upswing." *Journal of Banking and Finance* 62 (January): 15–27.

Official Gazette of the Republic of Slovenia. 2005a. "Regulation on the Assessment of Credit Risk Losses of Banks and Savings Banks." Technical Report No. 67a.

———. 2005b. "Regulation on the Assessment of Credit Risk Losses of Banks and Savings Banks." Technical Report No. 67b.

———. 2015. "Regulation on the Assessment of Credit Risk Losses of Banks and Savings Banks." Technical Report No. 50.

Oster, E. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business and Economic Statistics* 37 (2): 187–204.

Peek, J., and E. S. Rosengren. 2005. "Unnatural Selection: Perverse Incentives and the Misallocation of Credit in Japan." *American Economic Review* 95 (4): 1144–66.

Rubio, M., and J. A. Carrasco-Gallego. 2016. "The New Financial Regulation in Basel III and Monetary Policy: A Macroprudential Approach." *Journal of Financial Stability* 26 (October): 294–305.

Schuler, M. S., W. Chu, and D. Coffman. 2016. "Propensity Score Weighting for a Continuous Exposure with Multilevel Data." *Health Services and Outcomes Research Methodology* 16 (4): 271–92.

Tayler, W. J., and R. Zilberman. 2016. "Macroprudential Regulation, Credit Spreads and the Role of Monetary Policy." *Journal of Financial Stability* 26 (October): 144–58.

van Wijnbergen, S., and T. Homar. 2014. "On Zombie Banks and Recessions after Systemic Banking Crises: Government Intervention Matters."

Zhou, Q., C. McNeal, L. A. Copeland, J. P. Zachariah, and J. J. Song. 2020. "Bayesian Propensity Score Analysis for Clustered Observational Data." *Statistical Methods and Applications* 29 (2): 335–55.

Zimic, S. 2008. "Guverner marko kranjec je bankam 'priskrbel' 270 milijonov." ["Governor Marko Kranjec 'Provided' 270 Million Banks."] Dnevnik, October 22. https://www.dnevnik.si/1042216640.

The Cost of Breaking an Exchange Rate Peg: Synthetic Control Estimation*

Jae Hoon Choi[a] and A. Christopher Limnios[b]
[a]Xavier University
[b]Providence College

This paper uses synthetic control estimation to measure the impact of switching exchange rate regimes from fixed to floating on sovereign credit risk. The study confirms the statistically significant short-term cost of breaking a peg. The results demonstrate that breaking the exchange rate peg incurs an increase in the average risk premium by 0.22–0.34 standard deviations for two to five months. The study also investigates peg-formation episodes and finds that switching the regime from floating to fixed and switching the regime from fixed to floating have asymmetric impacts on the country risk spread, and it confirms the hypothesis that investors consider breaking an exchange rate peg as breaking central bank's commitment to monetary stability.

JEL Codes: F3, F31, F41.

1. Introduction

In recent years, more developing countries have switched their exchange rate regimes from the fixed to the more flexible ones to motivate economic growth (Batini et al. 2006; Klein and Shambaugh 2008). Since the choice of exchange rate regime is one of the major

*We appreciate Michael Hutchison, Kenneth Kletzer, Carl Walsh, Johanna Francis, Javier Gardeazabal, and Eric Aldrich for their valuable insights and comments. We thank all discussants and participants of the Western Economics Association International 94th Annual Meeting, the Southern Economics Association 89th Annual Meeting, the University of California Santa Cruz Macro workshop, the Xavier University Brown Bag seminar, and the Providence College Economics Department "Early Works in Progress" seminar for valuable feedback on this work. We also thank Linda Goldberg (editor) and two anonymous reviewers for their detailed and constructive comments as well as their efforts towards improving our manuscript.

international monetary policies constrained by the "trilemma" (Mundell 1963; Aizenman 2013; Popper, Mandilaras, and Bird 2013), it is crucial for central bankers to understand the benefits and costs of their exchange regime choice.

The current perspectives provided by the literature on the relationship between exchange rate regimes and economic welfare are contradictory. Levy-Yeyati and Sturzenegger (2003) find that fixed exchange rates are correlated with slower output growth rates and higher output volatility for the non-industrial countries. Jahjah, Wei, and Yue (2013) study how the choice of exchange rate regime affects the level of risk premium, and their study finds that the average country spread is 88 basis points lower in the countries under floating regimes than under fixed regimes. On the contrary, Alesina and Wagner (2006) find that switching exchange rate regimes from fixed to floating brings external costs into the economy. They investigate why countries employ different exchange rate regimes from their de jure regimes by studying the countries whose de jure and de facto exchange rate regimes do not match. The study concludes that breaking pegs may signal economic instability and that the financial market considers a wide exchange rate fluctuation as an indication of poor economic management. Their results show that switching exchange rate regimes from fixed to floating limits the effectiveness of interest rate or capital control policies. In addition, Aghion et al. (2009) find that exchange rate fluctuation may reduce firms' investment capacity and therefore lower the growth of the economy. On the other hand, while Ghosh, Gulde, and Wolf (2002) suggest a weak link between exchange rate regimes and real gross domestic product (GDP) growth, Harms and Kretschmann (2009) find that industrial countries benefit from flexible exchange rate fluctuations, and non-industrial countries may benefit more from implicit exchange rate stabilization.

Despite the vast literature on this subject, it is still challenging to quantify the trade-offs between regime choices because of the difficulty decoupling macroeconomic variables from the policy decisions of exchange rate intervention. The literature may benefit from a different approach to this subject. This paper investigates the impact of breaking the exchange rate peg on the country's risk premium using synthetic control estimation. More specifically, we investigate whether switching the exchange rate regime from fixed to floating in month T_0 leads to a higher country risk premium in the months T_i $(i > 0)$, compared with similar countries which keep the exchange rates pegged. In order to construct the synthetic control unit, comparison countries are selected and weighted by an algorithm based on their similarity to the treated country before the treatment with respect to both the outcome variable (country risk premium) and other covariates (macroeconomic indicators, such as consumer price index (CPI), trade volumes, and foreign reserves). Following Alesina and Wagner (2006), we hypothesize that foreign investors interpret the breaking of exchange rate pegs as the central banks' loss in the ability to keep their commitment to stabilizing their markets. As suggested in Billmeier and Nannicini (2013), the benefit of the synthetic control approach is that a linear combination of untreated countries allows a "transparent" estimation of the counterfactual outcome of the treated country. The "transparent" estimation of the costs of breaking a peg should show a clearer picture of international monetary policy decisions.

The results of this study suggest that abandoning the fixed exchange rate regime incurs increases in the country's risk premium. The increase in the risk premium can be interpreted as the additional cost of abandoning the fixed exchange rate regime because it implies that the price of foreign loans becomes more expensive for local entrepreneurs, and as entrepreneurs become more financially constrained, the growth of the economy lingers. However, we also find that such impacts may be short-lived because the statistical significance of the estimates of impacts do not last longer than two to five months.

The remainder of this paper is organized as follows. Section 2 describes the empirical strategy; Section 2.1 offers a brief description of the synthetic control analysis; Section 2.2 illustrates the basis of coding of de facto exchange rate regime that is adopted in this study, and Section 2.3 provides a summary description of the data set. Section 3 provides and discusses the results of synthetic control estimation and the statistical inferences using placebo tests in Section 3.1 and average treatment effect estimation results in Section 3.2. Sections 3.3, 3.4, and 3.5 present the extended analyses that check the robustness of the results. Section 3.6 discusses a potential extension of the study, and Section 4 concludes.

2. Empirical Strategy

2.1 Synthetic Control Method

Standard cross-country estimators on macro data tend to suffer multiple endogeneity issues, especially when one uses those models to estimate the effect of macroeconomic policies or government interventions (Billmeier and Nannicini 2013). The difference-in-differences method provides a quasi-experimental design to obtain a counterfactual to estimate a causal effect of an intervention. However, the difference-in-differences method avoids the critique of the endogeneity issues only under the parallel trend assumption, which is rarely plausible in international panel data. The synthetic control method, developed in Abadie and Gardeazabal (2003) and extended in Abadie, Diamond, and Hainmueller (2010, 2015), provides an alternative approach in comparative event studies to alleviate such potential concerns. This method applies a vector of weights to a subset of the total pool of candidate control countries, which keep their exchange rate fixed, creating a synthetic (counterfactual) country whose characteristics closely match the treated country's before the treatment. It then compares the risk premium index trajectory of the treated country, which abandons the peg, with the estimated risk premium index of the synthetic control, which keeps the peg. It, therefore, captures the treatment effect of abandoning the policy of stable currency prices on the risk premium index.

We provide a description of the synthetic control method used in this study, following the notation of Abadie, Diamond, and Hainmueller (2010, 2015).[1] Of the $J + 1$ units (countries), the first unit switches its exchange rate regime from fixed to floating, while the J others (the units $j = 2$ to $j = J + 1$) keep their exchange rate fixed and are referred to as the control country candidates. For countries $j = 1, \ldots, J + 1$ and months $t = 1, \ldots, T$, let $T_0 - 1$ be the number of pre-breaking-peg months, with $1 < T_0 < T$, and we assume that the sample is a balanced panel. This study constructs a synthetic control using the 10 months of data before the month that

[1] For technical details of the methodology, see Appendix B of Abadie, Diamond, and Hainmueller (2010) and footnote 5 of Abadie, Diamond, and Hainmueller (2015).

the country abandons the exchange rate peg.[2] Then, we investigate the synthetic control's risk premium and its difference from the peg-breaking country's risk premium for 10 months after the peg breaks. That is, we set $T_0 = 11$ and $T = 21$.

We define a synthetic control as a weighted average of the units in the control candidates (the "donor pool"); a synthetic control can be represented by a $(J \times 1)$ vector of weights $W = (w_2, \ldots, w_{J+1})'$, with $0 \leq w_j \leq 1$ for $j = 2, \ldots, J+1$ and $\sum_{j=2}^{J+1} w_j = 1$. Let X_1 be a $(k \times 1)$ vector containing the values of treated country's characteristics before it breaks the peg, and let X_0 be the $k \times J$ matrix with the values of the same variables for the control country candidates.[3] The elements in X_1 and X_0 in this study are the monthly indicators from the Global Economic Monitor (GEM, the World Bank) to capture the conditions and characteristics of each country.[4] Then, the gap between values of characteristics variables of the treated country and the synthetic control is expressed as $X_1 - X_0 W$, and we select W^*, which minimizes this gap for each episode of abandoning exchange rate peg. In terms of weight selection, Abadie, Diamond, and Hainmueller (2015) suggest the following process. For $m = 1, \ldots, k$, let X_{1m} be the value of the m-th variable for the peg-breaking country and X_{0m} be a $1 \times J$ vector containing the values of the m-th variable for the control country candidates. Then, we choose W^* as the value of W that minimizes the root mean square prediction error (RMSPE),

$$\sum_{m=1}^{k} \left(X_{1m} - X_{0m} W\right)^2 . \tag{1}$$

Then, let Y_{jt} be the outcome variable (the standardized country risk spread in this study) of country j at time t, and let Y_1 be a $(T_1 \times 1)$ vector of the post-breaking-peg values of the outcome

[2] De facto fixed exchange rate regime periods are identified using 10 months of consecutive low volatility in exchange rates; a detailed description is presented in Section 2.2.

[3] The pre-intervention characteristics in X_1 and X_0 may include pre-intervention values of the outcome variable (Abadie, Diamond, and Hainmueller 2010, 2015).

[4] A detailed description of the characteristics variables used in this study is presented in Section 2.3.

for the treated country: $Y_1 = (Y_{1\ T_0}, \ldots, Y_{1\ T})'$. Similarly, let Y_0 be a $(T_1 \times J)$ matrix, where column j contains the post-breaking-peg values of the outcome variable for unit $j + 1$. Then, the effect of switching regime (α_{1t}) for months $t \geq T_0$ is estimated as the difference between the predicted outcome of the synthetic control country and what is actually observed,

$$\hat{\alpha}_{1t} = Y_{1\ t} - \sum_{j=2}^{J+1} w_j^* Y_{j\ t}. \tag{2}$$

This study examines the 19 episodes of abandoning exchange rate pegs, which provide 19 series of $\hat{\alpha}_{1t}$ for $t \in [T_0,\ T]$. Then, we provide the average treatment effect estimates with statistical inferences whether the treatment effect is statistically different from zero for each t.

2.2 Identification of Regime-Switching Events

We identify monthly foreign exchange rate regime switches using a modified methodology of Klein and Shambaugh (2008). We classify a nominal exchange rate as "managed" if the month-end official bilateral exchange rate stays within a 2 percent band for 10 months, while pegs that last less than 10 months are classified as floating. Then, we confirm that these events align with the same regime switches specified in Klein and Shambaugh (2008);[5] our study uses this modified coding for identification at a monthly frequency, while Klein and Shambaugh (2008) provide regime classification at an annual frequency.

In order to check the robustness of results, our study also presents the results of investigating the episodes identified with a wider (2.5 percent) band in Section 3.3. This alternative method secures a larger number of control country candidates, while maintaining the statistical significance of de facto regimes. The quality of synthetic controls in terms of imitating the treated countries may depend on the size of the pool of control country candidates; a larger number of control candidates potentially raises the likelihood of finding a

[5] An updated exchange rate regime specification for 1960–2018 is provided by the author at https://iiep.gwu.edu/jay-c-shambaugh/data/.

better combination of control countries to construct the synthetic country that represents the treated country. Therefore, with a slight increase in the size of the band, we secure more control countries whose exchange rate regimes are fixed/managed during each study period and test whether the results are robust. One may argue that a country/period may be misclassified as fixed/managed simply due to a lack of shocks. However, according to Calvo and Reinhart (2002), the probability of having exchange rate volatility less than 2.5 percent within a month while having flexible exchange rates is between 60 percent and 70 percent. Therefore, the probability of having 10 consecutive months of exchange rate growth within a 2.5 percent band is estimated between 0.6 percent and 3 percent, which is still statistically insignificant.

2.3 Data

For the measure of sovereign credit risk, we use the J.P. Morgan Emerging Market Bond Index Global (EMBIG), a major index of emerging market economy sovereign bond spreads. The index is created as a benchmark reflecting returns from price gains and interest income on a "passive" portfolio of traded emerging markets debt. It is constructed as a composite of four markets: Brady bonds, Eurobonds, U.S. dollar local markets, and loans.

The index reflects country risk but is independent of exchange risk because the index presents the spread between the dollar-denominated bonds and the U.S. Treasury bill. For this reason, the index is often used in the sovereign risk literature (Kaminsky and Schmukler 2002; Gapen et al. 2008; Remolona, Scatigna, and Wu 2008; Hilscher and Nosbusch 2010), and it fits well for the purpose of this study; we intend to extract investors' sentiment from a country's exchange rate policy change on its sovereign risk, and it requires a measure that is highly correlated with sovereign risk and independent of (or less correlated with) exchange risk.[6] In other words, when central banks break exchange rate pegs by loosening (or

[6]In Section 3.5, we extend the study to investigate the episodes of peg formation and confirm the validity of the index for this study; the risk premium index does not show a systemic change during the period of more stable exchange rates, and the results imply that the EMBIG index is essentially independent of exchange risks.

losing) control over local-currency prices, unless investors consider volatile exchange rates a signal of economic instability, the index of synthetic controls will show the same trend as the index of the treated.

The data include 42 countries, of which J.P. Morgan reports the monthly risk premium index (EMBIG) from January 1998 to December 2015. Table 1 reports the list of countries and the summary statistics of EMBIG by country. The summary statistics indicate that the difference in level and variance of EMBIG among countries in the sample can be considerable. See, for example, two Latin American countries, Argentina and Chile. Argentina's average EMBIG is 1,518, while Chile's is 149.4. The standard deviation of EMBIG of Argentina is 1,811.1, while for Chile, the standard deviation is 56.6. Such difference may not be negligible because it causes potential estimation biases towards the countries with higher levels and variance of EMBIG. Due to this potential bias, we standardize EMBIG as follows:

$$\text{Std.EMBIG}_{c,t} = \frac{\text{EMBIG}_{c,t} - \hat{\mu}(\text{EMBIG})_c}{\hat{\sigma}(\text{EMBIG})_c}, \tag{3}$$

where $\text{EMBIG}_{c,t}$ is the EMBIG observation of country c at time t, $\hat{\mu}(\text{EMBIG})_c$ is the sample average of EMBIG of country c, and $\hat{\sigma}(\text{EMBIG})_c$ is the sample standard deviation of EMBIG of country c. Standardizing the units allows the construction of the synthetic control without losing much explanatory power to unit adjustments. The standardized measure is also useful for statistical inference; in Section 3.2, we combine the estimated treatment effects for statistical analysis, and it is essential to match the first and second moments of data to avoid estimation biases.

To set up a synthetic control framework, the following information is required for the country of interest as well as the control countries to construct its synthetic control: (i) a full set of EMBIG and (ii) a full set of predictor variables during the study period (10 months before and after the events). In addition, we exclude cases of economic (financial and currency) crisis; for example, we exclude the peg-breaking episodes during global crises, such as the 2007–09 subprime mortgage crisis, as well as during local crises, such as the

Table 1. Summary Statistics of EMBIG from January 1998 to December 2015

Country	Mean	SD	Min.	Max.
Algeria	826	424	304.4	2,306
Argentina	1,518	1,811.1	202.5	6,847
Belarus	717.3	284.8	266.7	1,582.3
Belize	1,028.3	454	383.4	2,427.7
Brazil	478.1	371.7	143.3	2,057.4
Bulgaria	341.9	276.7	43.1	1,365.9
Chile	149.4	56.6	54.9	383.1
Colombia	337.2	199.6	108.4	986
Côte d'Ivoire	1,551.7	1,050.3	297	3,407.8
Croatia	303.8	161.2	97.9	881.8
Dominican Rep.	492.4	299.6	138.9	1,709
Ecuador	1,146.7	844.9	354.4	4,416.2
Egypt	297.9	173.9	29.1	678.7
El Salvador	374	140.8	115.5	888.9
Gabon	461.3	213	225.7	1,215.3
Georgia	487.7	336.6	167.6	1,883.3
Ghana	591.6	249.2	321.1	1,572
Greece	103.3	22	57.3	152.6
Hungary	177.7	142.5	13.9	650.2
Indonesia	271.1	126.1	21.2	890.8
Iraq	582.1	202.1	20.2	1,239.4
Jamaica	522.3	175.1	276.6	1,120.8
Jordan	371	62.2	182.3	519.1
Kazakhstan	385.4	228.3	176.3	1,302.4
Korea	221.8	144.3	77.3	800.7
Malaysia	182.9	120	67.8	1,033.2
Mexico	268.3	138.7	97.6	944.1
Morocco	317.6	193.5	54.1	1,140.1
Nigeria	726.3	551.4	23.1	2,562.6
Pakistan	633.2	416.1	137.5	2,136.6
Panama	272	124.6	102.2	603.7
Peru	312.1	195.4	104	935.8
Philippines	298	169.7	82.8	937.3
Poland	145.4	78.7	39.5	344.9
Senegal	466.8	106.6	267.6	720.2
Sri Lanka	515.3	355.1	256.4	2,092.6
Thailand	153.5	120.3	41.9	722.8
Turkey	393.1	208.7	162.4	1,048.3
Ukraine	825.2	708.4	140.3	3,863
Uruguay	351.3	278.2	121.4	1,501.8
Venezuela	1,244.1	926.2	186.6	4,892.9
Vietnam	282.1	139.7	9.1	999.8

1998–2002 Argentine great depression, the Venezuelan banking crisis of 2009–10, etc.[7] During financial and currency crises, it is likely that the country's risk premium index fluctuates not just because of an exchange rate regime switch but because of other factors.[8] Since financial crises often lead to currency crises and vice versa (Mishkin 1999), including these cases may result in an overestimation of the effect of switching regimes (or, rather, of being "forced" to abandon the fixed regime), although the predictor variables of the synthetic controls may be able to control for such crises to some extent.

As a robustness check, we also present the estimation results excluding political upheavals in Section 3.4; political instability, which raises the country risk premium, may accompany a loss of control over exchange rates, resulting in an overestimation of the impact of breaking a peg.

Based on the criteria described above, 19 peg-breaking cases are identified and studied in the synthetic control framework. Each peg-breaking episode is listed in Table 2 with the following information: the treated country, the peg-breaking year and month, and the control country candidates, which are the countries that keep their exchange rate pegs during the study period of each case and are used to construct the synthetic control of the treated country. One may question if it is necessary to consider geological proximity or cultural/political similarity between the treated country and its control country candidates; however, such requirements are unnecessary in synthetic control studies. First, the synthetic control method finds the best combination out of the pool of candidates using their covariates to construct the synthetic control country. If, for example, one of the control country candidates is in a different continent from the one where the treated country is located and if such geological distance is the factor that makes the risk premium of the control country candidate behave differently from one of the treated countries, the control country candidate will not be used (given zero or

[7]Table A.1 in Appendix A lists the peg-breaking episodes that are excluded from estimation due to coincidence with economic crises.

[8]For example, a country may undergo a corruption scandal which may increase uncertainty (or any other event which may trigger capital outflow), contributing to an increase in the risk premium. We control for these events by assuming that these types of crises almost always go hand in hand with financial/currency crises.

Table 2. Nineteen Peg-Breaking Episodes and Corresponding Donor Pools

Country	Year	Month	Control Country Candidates
Peru	2001	July	Malaysia, Panama
Egypt	2005	February	Ecuador, El Salvador, Malaysia, Pakistan, Panama
Peru	2005	November	Ecuador, El Salvador, Malaysia, Pakistan, Panama, Philippines
Mexico	2006	April	Argentina, Ecuador, El Salvador, Pakistan, Panama, Philippines
Ecuador	2010	May	Belize, Dominican Republic, Egypt, El Salvador, Iraq, Panama, Sri Lanka, Vietnam
Ghana	2011	February	Argentina, Belize, Dominican Republic, Egypt, El Salvador, Iraq, Kazakhstan, Pakistan, Panama, Peru, Sri Lanka, Vietnam
Indonesia	2011	October	Argentina, Belize, Dominican Republic, Egypt, El Salvador, Iraq, Kazakhstan, Sri Lanka, Panama, Peru, Ukraine, Vietnam
Pakistan	2012	January	Argentina, Belize, Dominican Republic, Egypt, El Salvador, Iraq, Kazakhstan, Sri Lanka, Panama, Peru, Ukraine, Vietnam
Egypt	2013	February	Belize, Dominican Republic, Ecuador, El Salvador, Georgia, Iraq, Kazakhstan, Sri Lanka, Vietnam, Panama, Ukraine
Jamaica	2013	April	Belize, Dominican Republic, Ecuador, El Salvador, Iraq, Kazakhstan, Jordan, Peru, Ukraine, Uruguay
Argentina	2013	August	Belize, Dominican Republic, Ecuador, El Salvador, Iraq, Jordan, Panama, Sri Lanka, Vietnam
Indonesia	2013	September	Belarus, Belize, Dominican Republic, Ecuador, El Salvador, Iraq, Jordan, Panama, Sri Lanka, Vietnam
Georgia	2014	January	Belarus, Belize, Dominican Republic, Ecuador, El Salvador, Iraq, Panama, Sri Lanka, Vietnam
Kazakhstan	2014	March	Belarus, Belize, Dominican Republic, Ecuador, El Salvador, Iraq, Panama, Sri Lanka, Vietnam
Pakistan	2014	April	Belize, Dominican Republic, Ecuador, El Salvador, Iraq, Panama, Sri Lanka, Vietnam
Belarus	2015	February	Belize, Dominican Republic, Ecuador, El Salvador, Iraq, Panama, Sri Lanka, Venezuela, Vietnam
Egypt	2015	March	Belize, Dominican Republic, Ecuador, El Salvador, Iraq, Panama, Sri Lanka, Venezuela, Vietnam
Peru	2015	March	Belize, Dominican Republic, Ecuador, El Salvador, Iraq, Panama, Sri Lanka, Venezuela, Vietnam
Philippines	2015	September	Belize, Dominican Republic, Ecuador, El Salvador, Iraq, Panama, Sri Lanka, Vietnam

low weight) in synthetic control construction. Secondly, such filtering will limit the effectiveness of the method because it quickly lowers the number of control country candidates. For example, although there are 10 control country candidates in the case of Indonesia in September 2013, only 2 countries (Sri Lanka and Vietnam) are Asian countries.

The monthly indicators from the Global Economic Monitor (World Bank 2021a) are used as predictor variables to match and forecast the risk premium index. The indicators[9] include (i) consumer price index, (ii) import merchandise, (iii) export merchandise, (iv) stock market index, (v) months import cover of foreign reserves, (vi) industrial production, (vii) total reserves, and (viii) unemployment rate. Except for the unemployment rate, we use logged values for all variables.[10]

3. Results

In this section, we present and discuss the results of the experiments, highlighting specific cases. The results provide numerical

[9]The following data definition/description is from the Global Economic Monitor data set (World Bank 2021a): (i) Consumer price index: Data are in nominal terms and seasonally adjusted. (ii) Import merchandise: Merchandise (goods) imports, cost, insurance, and freight basis (c.i.f.), in constant US$ millions, seasonally adjusted. The base year is 2005. (iii) Export merchandise: Merchandise (goods) exports, free on board (f.o.b.), in constant US$ millions, seasonally adjusted. The base year is 2005. (iv) Stock market index: local equity market index in US$. (v) Months import cover of foreign reserves: The stock of international reserves is expressed as the number of months of financing coverage it represents for the given country's imports of merchandise goods. (vi) Industrial production: The output for the industrial sector of the economy. The industrial sector includes manufacturing, mining, and utilities. Data are in constant US$, seasonally adjusted. The base year is 2005. (vii) Total reserves: Total reserves comprise holdings of monetary gold, special drawing rights, reserves of International Monetary Fund members held by the IMF, and holdings of foreign exchange under the control of monetary authorities. Data are in constant US$ millions. The base year is 2005. (viii) Unemployment rate: Data are in percent.

[10]One may notice that the imports, exports, industrial production, and total reserves are not scaled with the country's GDP or monetary base, which are not available for all countries at monthly frequency. Scaling the variables is not necessary for the prediction process in the synthetic control methodology because the synthetic control algorithm solves for a set of optimal country-specific weights using the combination of as well as controlling for available indicators that are highly correlated with GDP.

and graphical evidence that breaking a peg, in general, increases the country's risk premium and eventually induces additional costs that monetary authorities should consider.

We first present, in Table 3, the weights of the control countries selected to construct the synthetic control country in each case. As explained in Section 2.1, the control countries are weighted based on their similarity (by minimizing RMSPE) to the treated country before the treatment; the weighting algorithm uses both the outcome variable (country risk premium) and other covariates (macroeconomic indicators of GEM). For example, in the case of Ecuador in May 2010, the trend of the risk premium in Ecuador, before switching the exchange regime from fixed to floating, is best reproduced by a combination of the risk premium trends and covariates of the Dominican Republic (13.2 percent), Egypt (2.8 percent), Pakistan (17.9 percent), Panama (18.1 percent), and Vietnam (65.8 percent). Other control candidates (Belize, El Salvador, and Sri Lanka) are assigned zero weights.[11]

In Table 4, we present the numerical comparison of covariates between each treated country and the constructed synthetic control; the table reports the average values of predictors of the treated unit and the synthetic control unit. They show how closely the synthetic control country captures the characteristics of the treated country. Except for one case[12] of an unbalanced predictor, the average values of predictors of treated units and synthetic control units are overall balanced, and it indicates that the weights are well calibrated to construct the synthetic units that imitate the treated units.

Figure 1 reports the trends of standardized EMBIG of treated units (navy solid line) and synthetic control units (red dashed line) for all cases. The trends of the synthetic EMBIG (the counterfactual) closely track the trajectories of the trend of the "actual" EMBIG until the regime switches; during the period prior to the regime switch, the weights for the predictors from GEM of control candidates are calibrated to best match the trajectory of the treated

[11] The weights are selected through the permutation process of the synthetic control methodology. Therefore, the number of control countries receiving non-zero weights differs by case; for example, all nine control country candidates get non-zero weights in the case of Georgia (2014), while only one (of nine) control country candidate gets non-zero weights in the case of the Philippines (2015).

[12] 1 of 65 predictors: Months Import Cover of Mexico (April 2006).

Table 3. Country Weights in Synthetic Countries

Case	Controls	Weight	Case	Controls	Weight	Case	Controls	Weight	Case	Controls	Weight
PER	MYS	100%	**GHA**	ARG	0%	**ARG**	BLZ	0%	**BLR**	BLZ	0%
Feb. 2005	PAN	0%	Feb. 2011	BLZ	0%	Aug. 2013	DOM	0%	Feb. 2015	DOM	0%
				DOM	15.3%		ECU	27.8%		ECU	49.7%
				EGY	12.1%		IRQ	6.9%		IRQ	0%
				SLV	0%		JOR	16.8%		LKA	0%
				IRQ	0%		LKA	0%		PAN	0%
				KAZ	0%		PAN	0%		SLV	0%
				LKA	72.7%		SLV	48.5%		VEN	50.3%
				PAK	0%		VNM	0%		VNM	0%
				PAN	0%						
				PER	0%						
				VNM	0%						
EGY	ECU	0%	**IDN**	ARG	85.8%	**IDN**	BLR	1.3%	**EGY**	BLZ	0%
Feb. 2005	SLV	0%	Oct. 2011	BLZ	0%	Sep. 2013	DOM	0%	Mar. 2015	DOM	0%
	MYS	0%		DOM	0%		ECU	0%		ECU	0%
	PAK	100%		EGY	0%		IRQ	80.6%		IRQ	41.1%
	PAN	0%		IRQ	0%		JOR	0%		LKA	0%
				KAZ	0%		LKA	17.2%		PAN	14.2%
				LKA	0%		PAN	0.9%		SLV	44.7%
				PAN	0%		SLV	0%		VEN	0%
				PER	0%		VNM	0%		VNM	0%
				SLV	0%						
				UKR	0%						
				VNM	14.2%						
PER	ECU	0%	**PAK**	ARG	0%	**GEO**	BLR	0.1%	**PER**	BLZ	0%
Nov. 2005	SLV	23.8%	Jan. 2012	BLZ	1.0%	Jan. 2014	BLZ	35.9%	Mar. 2015	DOM	0%
	MYS	0%		DOM	0%		DOM	0.9%		ECU	10.3%
	PAK	47.0%		EGY	49.6%		ECU	3.8%		IRQ	0%
	PAN	0%		IRQ	0%		IRQ	24.1%		LKA	0%
	PHL	29.2%		KAZ	0%		LKA	33.7%		PAN	0%
				LKA	0%		PAN	0.3%		SYL	0%
				PAN	0%		SYL	0.6%		VEN	89.7%
				PER	0%		VNM	0.6%		VNM	0%
				SLV	0%						
				UKR	49.4%						
				VNM	0%						

(continued)

Table 3. (Continued)

Case	Controls	Weight	Case	Controls	Weight	Case	Controls	Weight	Case	Controls	Weight
MEX	ARG	0%	**EGY**	BLZ	12.9%	**KAZ**	BLR	7.7%	**PHL**	BLZ	0%
Apr. 2006	ECU	0%	Feb. 2013	DOM	0%	Mar. 2014	BLZ	0%	Sep. 2015	DOM	0%
	SLV	0%		ECU	0%		DOM	17.5%		ECU	0%
	PAK	82.1%		GEO	0%		ECU	55.9%		IRQ	0%
	PAN	17.9%		IRQ	0%		IRQ	18.8%		LKA	100%
	PHL	0%		KAZ	0%		LKA	0%		PAK	0%
				LKA	0%		PAN	0%		PAN	0%
				PAN	0%		SLV	0%		SLV	0%
				SLV	0%		VNM	0%		VNM	0%
				UKR	87.1%						
				VNM	0%						
ECU	BLZ	0%	**JAM**	BLZ	8.8%	**PAK**	BLZ	2.3%			
May 2010	DOM	13.2%	Apr. 2013	DOM	0%	Apr. 2014	DOM	51.8%			
	EGY	2.8%		ECU	0%		ECU	0%			
	SLV	0%		IRQ	0%		IRQ	0%			
	PAK	17.9%		KAZ	87.8%		LKA	0%			
	PAN	18.1%		JOR	0%		PAN	39.0%			
	LKA	0%		LKA	0%		SLV	0%			
	VNM	65.8%		PAN	0%		VNM	6.9%			
				SLV	0%						
				UKR	3.4%						
				VNM	0%						

Note: ISO Alpha-3 country codes are used to indicate country names in this table: Algeria (DZA), Argentina (ARG), Belarus (BLR), Belize (BLZ), Brazil (BRA), Bulgaria (BGR), Chile (CHL), Colombia (COL), Côte d'Ivoire (CIV), Croatia (HRV), Dominican Republic (DOM), Ecuador (ECU), Egypt (EGY), El Salvador (SLV), Gabon (GAB), Georgia (GEO), Ghana (GHA), Greece (GRC), Hungary (HUN), Indonesia (IDN), Iraq (IRQ), Jamaica (JAM), Jordan (JOR), Kazakhstan (KAZ), Korea (KOR), Malaysia (MYS), Mexico (MEX), Morocco (MAR), Nigeria (NGA), Pakistan (PAK), Panama (PAN), Peru (PER), Philippines (PHL), Poland (POL), Senegal (SEN), Sri Lanka (LKA), Thailand (THA), Turkey (TUR), Ukraine (UKR), Uruguay (URY), Venezuela (VEN), Vietnam (VNM).

Table 4. Predictor Balance in Synthetic Country Construction

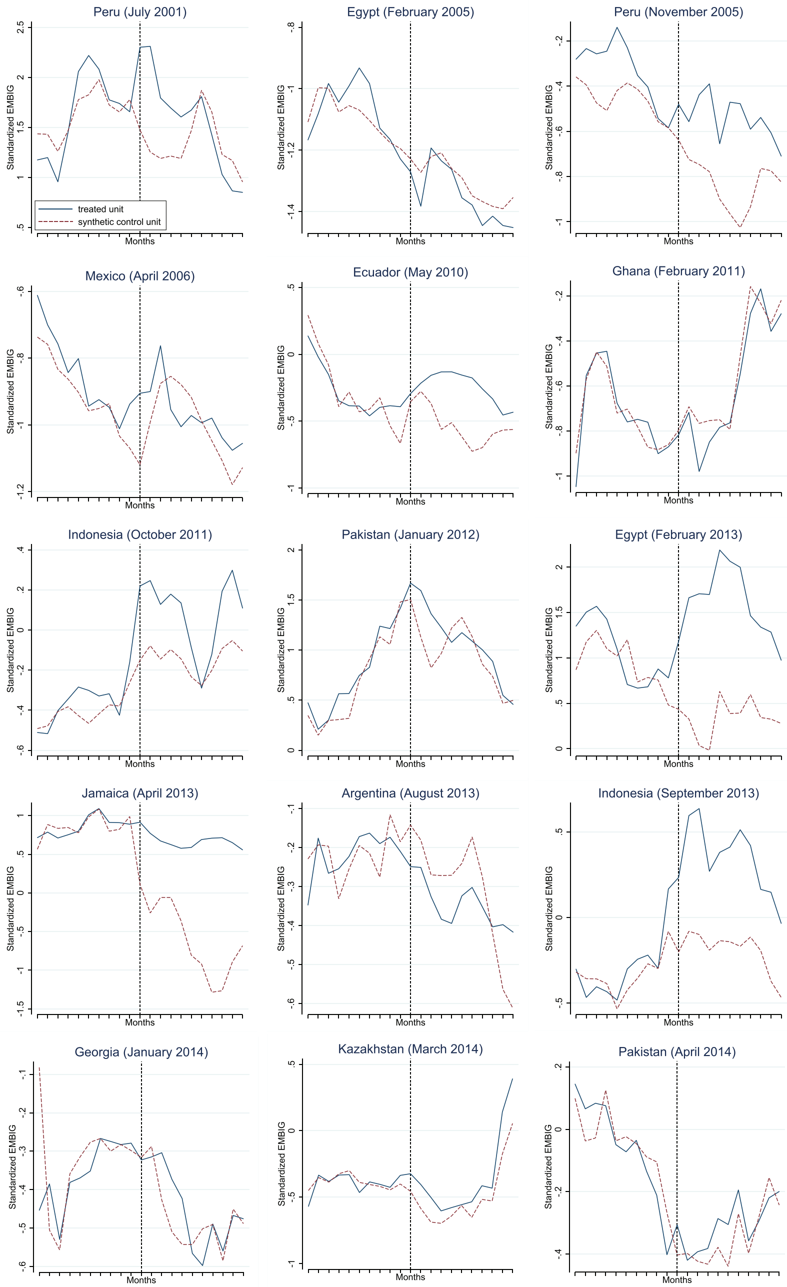
	Treated	Synthetic		Treated	Synthetic		Treated	Synthetic
Peru, Jul. 2001			**Pakistan, Jan. 2012**			**Pakistan, Apr. 2014**		
Imports Merchandise	6.929	5.859	Exports Merchandise	7.624	8.392	Consumer Price Index	4.972	4.867
Exports Merchandise	6.367	7.320	Months Import Cover	1.589	0.993	Exports Merchandise	7.666	7.666
Months Import Cover	2.654	1.225	Total Reserves	9.805	9.795	Total Reserves	9.013	9.013
Total Reserves	9.085	6.829	**RMSPE**		0.112	**RMSPE**		0.079
RMSPE		0.265						
Egypt, Feb. 2005			**Egypt, Feb. 2013**			**Belarus, Feb. 2015**		
Consumer Price Index	4.066	4.027	Consumer Price Index	4.822	4.873	Consumer Price Index	5.906	5.425
Imports Merchandise	7.258	7.564	Months Import Cover	0.907	0.801	Exports Merchandise	7.941	6.551
Exports Merchandise	6.499	7.002	Total Reserves	9.617	9.439	Total Reserves	8.323	9.345
Months Import Cover	2.561	2.047	**RMSPE**		0.297	**RMSPE**		0.374
Total Reserves	9.609	9.399						
RMSPE		0.238						
Peru, Nov. 2005			**Jamaica, Apr. 2013**			**Egypt, Mar. 2015**		
Consumer Price Index	4.474	4.247	Consumer Price Index	4.822	4.725	Consumer Price Index	5.025	4.773
Imports Merchandise	7.169	7.651	Exports Merchandise	4.995	6.360	Exports Merchandise	7.636	7.638
Exports Merchandise	7.218	7.084	Months Import Cover	0.763	1.701	Total Reserves	9.702	9.619
Months Import Cover	2.635	1.551	Total Reserves	7.063	8.992	**RMSPE**		0.228
Total Reserves	9.523	9.037	**RMSPE**		0.094			
RMSPE		0.162						

(*continued*)

Table 4. (Continued)

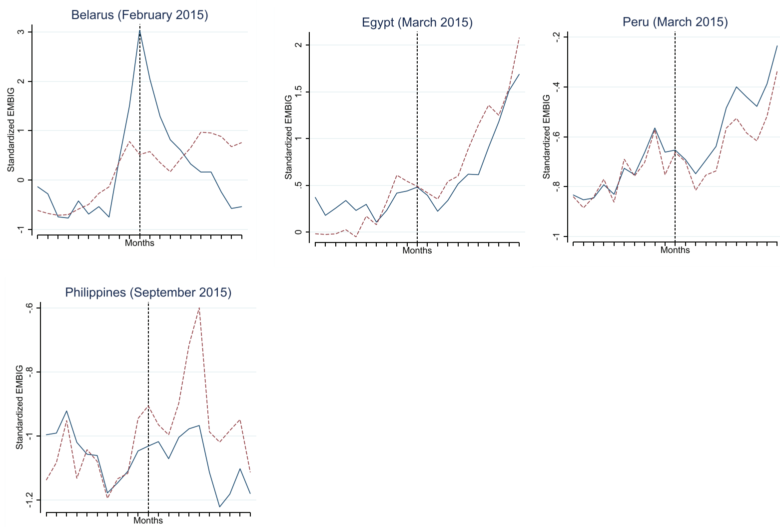
	Treated	Synthetic		Treated	Synthetic		Treated	Synthetic
Mexico, Apr. 2006			**Argentina, Aug. 2013**			**Peru, Mar. 2015**		
Imports Merchandise	10.022	8.000	Exports Merchandise	8.810	8.129	Consumer Price Index	4.750	4.799
Exports Merchandise	9.858	7.392	Months Import Cover	1.897	1.739	Exports Merchandise	8.058	7.344
Months Import Cover	–1.169	1.585	Total Reserves	10.620	9.637	Total Reserves	11.293	8.308
Total Reserves	8.708	9.430	**RMSPE**		0.064	**RMSPE**		0.062
RMSPE		0.048						
Ecuador, May 2010			**Indonesia, Sept. 2013**			**Philippines, Sept. 2015**		
Consumer Price Index	4.596	4.592	Consumer Price Index	4.763	4.763	Consumer Price Index	4.778	4.801
Exports Merchandise	7.221	7.365	Imports Merchandise	9.669	8.123	Exports Merchandise	8.514	7.216
Months Import Cover	1.068	1.241	Exports Merchandise	9.628	8.642	Total Reserves	9.628	8.777
Total Reserves	8.314	8.305	Months Import Cover	1.891	2.363	**RMSPE**		0.426
RMSPE		0.126	Total Reserves	11.559	10.506			
			RMSPE		0.250			
Ghana, Feb. 2011			**Georgia, Jan. 2014**					
Exports Merchandise	6.574	7.138	Consumer Price Index	4.722	4.722			
Months Import Cover	1.255	1.277	Total Reserves	8.010	8.009			
Total Reserves	8.220	8.229	**RMSPE**		0.128			
RMSPE		0.068						
Indonesia, Oct. 2011			**Kazakhstan, Mar. 2014**					
Exports Merchandise	9.723	8.825	Consumer Price Index	4.833	4.828			
Months Import Cover	2.038	1.916	Exports Merchandise	8.867	7.715			
Total Reserves	11.592	10.665	Total Reserves	10.137	8.871			
RMSPE		0.071	**RMSPE**		0.058			

Figure 1. Trends of Standardized EMBIG of Treated Country (solid line) vs. Synthetic Control (dashed line)



(*continued*)

Figure 1. (Continued)



country's risk premium by minimizing RMSPE. The estimated cost of abandoning stable exchange rates ($\hat{\alpha}_t$) is represented by the difference between the standardized EMBIG of the treated unit and the synthetic control unit after switching its exchange rate regime from fixed to floating.

Overall, the results support the hypothesis that abandoning a regime of stable exchange rates increases the country's risk premium; in other words, the cost of breaking a peg exists and is estimated to be positive. In most cases, the standardized EMBIG of the treated countries (the solid lines) start deviating positively from the synthetic controls (the dashed lines) around the times when exchange rates start floating: Peru (2001), Peru (2005), Mexico (2006), Ecuador (2010), Indonesia (2011), Pakistan (2012), Egypt (2013), Jamaica (2013), Indonesia (2013), Georgia (2014), Kazakhstan (2014), Pakistan (2014), Belarus (2015), and Peru (2015). These cases suggest that if the country kept its exchange rate fixed, it would have maintained a lower risk premium compared with the actual risk premium. In two cases, Egypt (2015) and the Philippines (2015), the treated country's risk premium decreases more than the synthetic control country's risk premium. In three cases—Egypt

(2005), Ghana (2011), and Argentina (2013)—switching regimes does not seem to affect sovereign risk because the trajectories of solid and dashed lines do not deviate from each other in the sample periods.

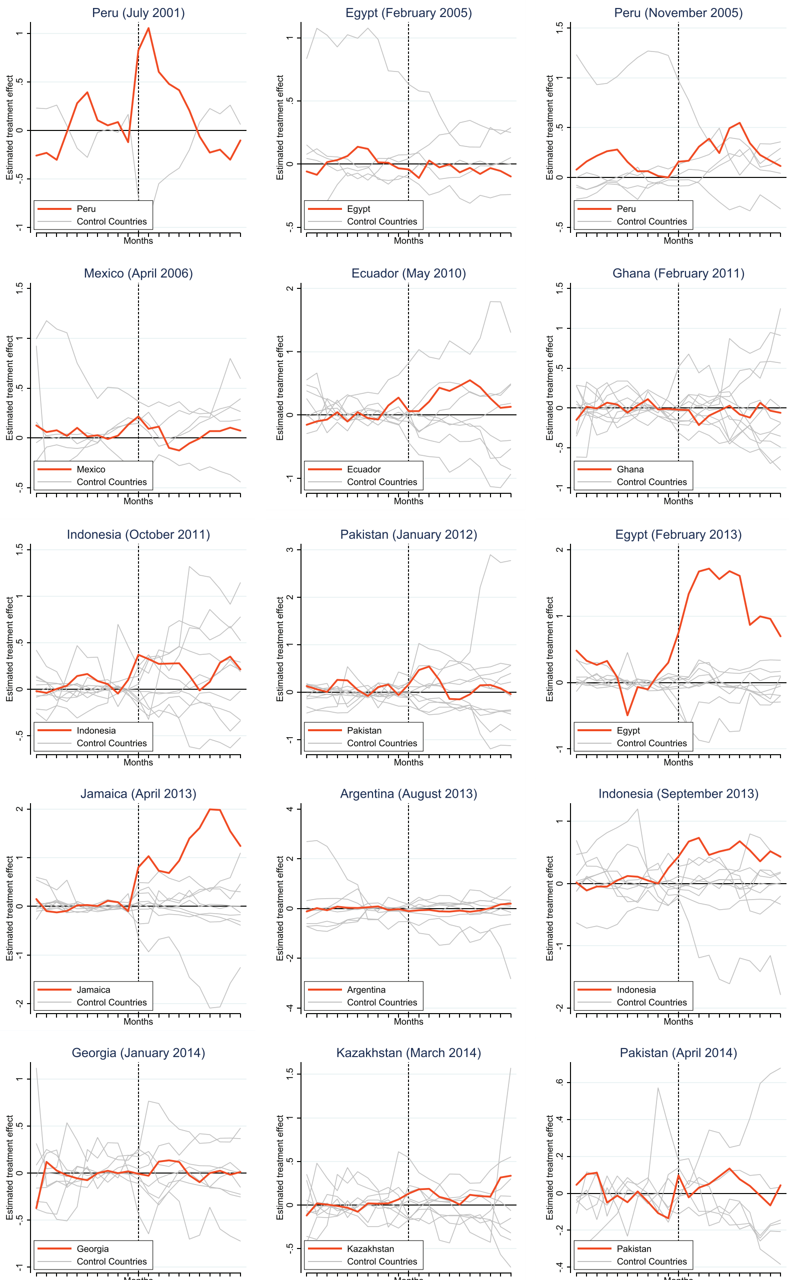
In some cases, the deviation between a treated unit and a synthetic control unit seems to start before (one to two months) the regime switch occurs: for example, Ecuador (2010), Egypt (2013), Indonesia (2013), and Belarus (2015). This can be explained by the limitation of investigating monthly averages of higher-frequency observations. Exchange rates are updated every second of every day, and their monthly average is the average of these values within a month. If the central bank switches its regime from fixed to floating in the second half of January, the exchange rates in January are averaged out, and currency price volatility may show up only in February, even though the regime was switched in January. Therefore, the risk premium index seems to deviate earlier than expected because the market reacts to higher-frequency information.

3.1 Inference from Placebo Tests

Following the literature (Abadie and Gardeazabal 2003; Abadie, Diamond, and Hainmueller 2010), we evaluate the statistical significance of our estimates through placebo tests. The placebo test demonstrates that the treatment effect is statistically not different from zero when it "should not" exist. The placebo studies are performed by applying the synthetic control method to the control countries. Since the countries in the control group did not switch their regimes, we "should not" observe any systematic change in the trends of their risk premium index. We expect the following if the synthetic control method successfully captures the effect of switching regimes in the treated countries: (i) the estimated α_{1t} of both the treated and control country stay at zero before switching regimes, and (ii) after the exchange rate regime is switched in the treated country, the estimated α_{1t} of the treated country shows a sudden increase, while the estimated α_{1t} of the control countries should remain statistically equivalent to zero or follow a random walk.

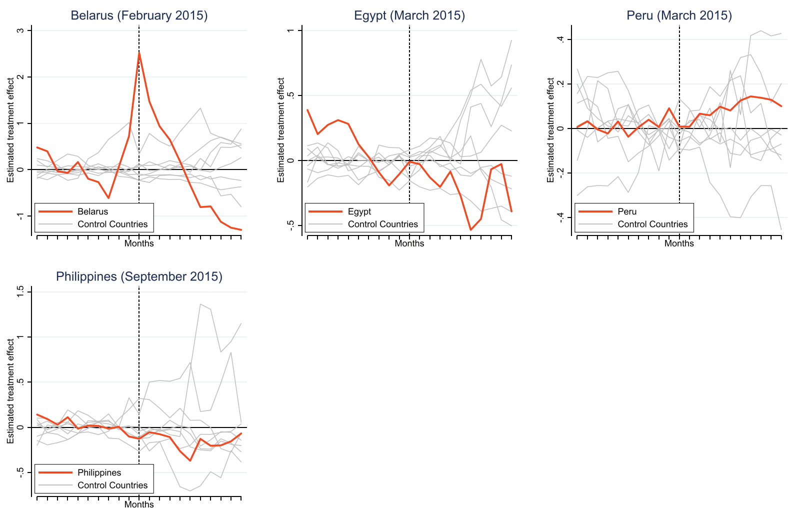
Figure 2 provides the placebo test results, and they support our hypothesis. We present the trajectory of the estimated α_{1t} for the treated country which switched its regime from fixed to floating

Figure 2. Trends of Estimated Treatment Effect (orange line) vs. Placebo Effect (gray line)



(continued)

Figure 2. (Continued)



(orange/darker line) and the trajectories of the control countries which maintained their fixed regimes (gray/lighter line). Most of them show that the treated countries' estimated α_{1t} start deviating from zero or follow a hump-shaped pattern after the events, while the ones for control countries do not follow this stylized pattern. These patternless gray lines indicate that the synthetic control estimation in this study does not systematically create or capture a false treatment effect. However, it is also worth noting that the placebo test does not directly infer that the treatment effect on the treated country is statistically different from zero, either. In the following section, we address this issue using a classical statistical inference method.

3.2 Inference from Average Treatment Effect Estimates

Despite its clear intuition, questions have been raised on the reliability of statistical inference in placebo tests. Most synthetic control studies, including Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2010), examine the treatment effect on one treated unit. Since it creates an estimated trend of time series for only one unit, there are not enough samples to provide statistical significance. Therefore, the best way to show this effect in such studies is through permutation exercises, which demonstrate that this effect is not observable in control units, that is, placebo tests. However, several studies argue that placebo tests in synthetic control analyses may be misleading. For example, Hahn and Shi (2017) show by using a data-generating process (DGP) that placebo tests may fail to provide proper statistical inference when a pre-estimation is made on the observations with size distortions. On the other hand, Ferman and Pinto (2017) present examples that placebo tests can have size distortions even when they consider an infeasible synthetic control estimator that correctly reconstructs the factor loadings of the treated unit.

Therefore, in order to avoid the potential issues in placebo tests discussed above and complement our findings, we present "classical" estimates of the treatment effect as well. Specifically, we find the mean of treatment effect estimates for a given t of all cases:

$$\hat{\hat{\alpha}}_{1t} = \frac{1}{n}\sum_{i=1}^{n} \hat{\alpha}_{1it}, \tag{4}$$

where $\hat{\alpha}_{1it}$ denotes the treatment effect estimate of country i at time t and n denotes the number of cases. Since we examine multiple cases ($n = 19$) of treatment effect and also use a standardized measure of EMBIG—as shown in Equation (3)—for all treated and control countries, combining these results allows us to estimate the average treatment effect and obtain the standard errors of the estimates without estimation bias from the significant differences in the first and second moments of the index among countries. The average effect of switching regimes is summarized in Figure 3 and Table 5. The estimates show that the effect of breaking an exchange rate peg on a country's risk premium is positive and statistically significant. Figure 3 shows that the treatment effect estimates are statistically significant for two to five months after switching its exchange rate regime (two months with 99 percent confidence intervals and five months with 95 percent confidence intervals); in magnitude, EMBIG rises by 0.3203–0.3439 standard deviations with 99 percent confidence intervals (0.2164–0.3439 standard deviations with 95 percent confidence intervals).

Figure 3. Average Treatment Effect Estimates $\hat{\hat{\alpha}}_{1t}$ and Statistical Significance Over Time

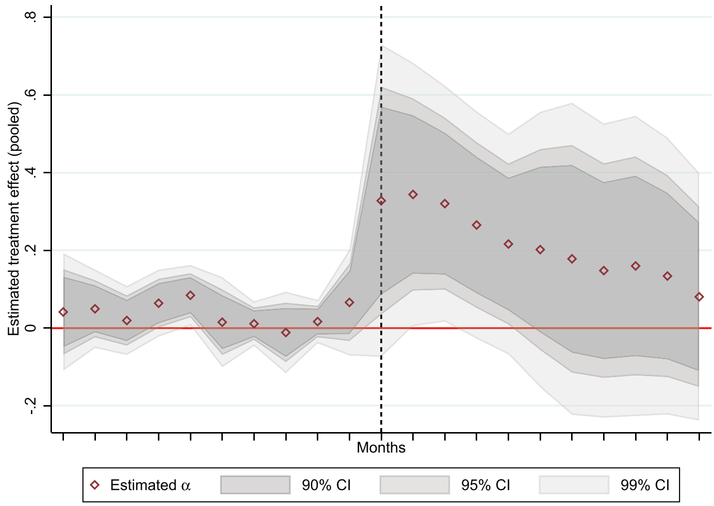


Table 5. Estimated Treatment Effect and Statistical Significance Over Time (p-values are calculated for the hypotheses set of $H_0 : \alpha_{1t} = 0$ and $H_1 : \alpha_{1t} \neq 0$)

Before Switching			After Switching		
t	$\hat{\hat{\alpha}}_{1t}$	p-value	t	$\hat{\hat{\alpha}}_{1t}$	p-value
1	0.0415	0.4403	T_0	0.3278**	0.0307
2	0.0497	0.1746	T_1	0.3439***	0.0092
3	0.0196	0.5355	T_2	0.3203***	0.0072
4	0.0641**	0.0471	T_3	0.2654**	0.0178
5	0.0845***	0.0060	T_4	0.2164**	0.0419
6	0.0152	0.7107	T_5	0.2020	0.1186
7	0.0113	0.5811	T_6	0.1782	0.2180
8	−0.0113	0.7602	T_7	0.1480	0.2755
9	0.0168	0.4007	T_8	0.1599	0.2490
10	0.0662	0.1808	T_9	0.1340	0.2940
			T	0.0806	0.4764

Note: *** indicates 99 percent significance level; ** indicates 95 percent significance level; * indicates 90 percent significance level.

3.3 Robustness Check: Alternative de facto Fixed Regime Specification

In this section, we use a slightly wider band (2.5 percent) than the one used in Klein and Shambaugh (2008) to identify the events and treated/control countries. As explained in Section 2.2, the 2.5 percent bands potentially secure a larger pool of control units while keeping the validity of the band, so the probability of having 10 consecutive months of spurious pegs within the band becomes statistically insignificant.

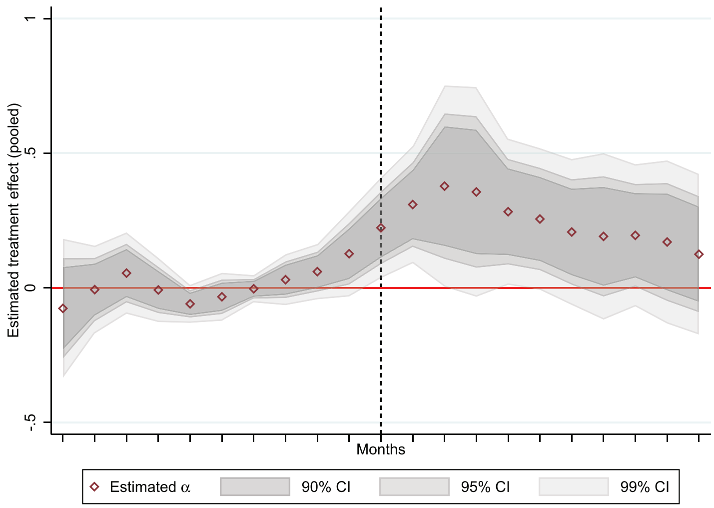
Table 6 lists the 16 identified cases and corresponding control country candidates. 2.5 percent bands identify fewer peg-breaking cases but allow more control country candidates per case. This is because while 2 percent bands identify more periods as de facto floating regime periods, they identify more periods as de facto fixed regime periods. This also explains why different peg-breaking events are identified under different sizes of bands, except for the cases of Mexico (2006), Pakistan (2012), and Indonesia (2013). For example, the case of Peru (2001) is considered a regime-switching case under a 2 percent band but not with a 2.5 percent band, and the case of Korea (2003) is identified with a 2.5 percent band but not with a 2 percent band; the Peru (2001) case is not identified as a regime-switching episode under a 2.5 percent band because the entire study period (September 2000–May 2002) is identified as a period of fixed regime due to the wider band, and the Korea (2003) case is not identified as a peg-breaking episode under a 2 percent band because the entire period (April 2002–December 2003) is classified as floating regime months due to the narrower band. Tables 14 and 15 in Appendix B report the weights of control countries to construct the synthetic control countries and the balance of covariates.

The analysis results using these episodes identified with the alternative de facto regime specification do not deviate much from the previous results and reconfirm the hypothesis that breaking the exchange rate peg has statistically significant impacts on the risk premium. The standardized EMBIG of the treated countries deviate positively from the synthetic controls after the exchange rate pegs are broken: Philippines (2000), Korea (2003), Argentina (2003), Bulgaria (2003), Hungary (2005), Poland (2005), Mexico (2006), Belarus (2011), Uruguay (2011), Mexico (2011), Pakistan (2012),

Table 6. Sixteen Peg-Breaking Episodes Identified with 2.5 Percent Bands

Country	Year	Month	Control Country Candidates
Mexico	1999	September	Argentina, Malaysia, Morocco, Panama, Peru, Venezuela
Philippines	2000	August	Argentina, Colombia, Ecuador, Malaysia, Panama, Peru
Korea	2003	February	Bulgaria, Côte d'Ivoire, Croatia, Ecuador, El Salvador, Hungary, Malaysia, Morocco, Philippines, Panama, Peru, Thailand, Ukraine, Poland
Argentina	2003	August	Colombia, Chile, Côte d'Ivoire, Croatia, Ecuador, El Salvador, Indonesia, Malaysia, Morocco, Philippines, Panama, Peru, Thailand, Ukraine
Bulgaria	2003	August	Côte d'Ivoire, Croatia, Ecuador, El Salvador, Malaysia, Morocco, Philippines, Panama, Peru, Thailand, Ukraine
Chile	2004	May	Côte d'Ivoire, Ecuador, Egypt, El Salvador, Hungary, Malaysia, Morocco, Philippines, Peru, Ukraine
Hungary	2005	February	Ecuador, Egypt, El Salvador, Malaysia, Mexico, Nigeria, Philippines, Peru, Ukraine, Uruguay
Poland	2005	May	Argentina, Colombia, Ecuador, Egypt, El Salvador, Indonesia, Malaysia, Mexico, Nigeria, Pakistan, Philippines, Peru, Ukraine, Uruguay
Mexico	2006	April	Argentina, Bulgaria, Côte d'Ivoire, Ecuador, Egypt, El Salvador, Malaysia, Morocco, Nigeria, Pakistan, Philippines, Peru, Ukraine, Uruguay, Venezuela
Belarus	2011	July	Argentina, Belize, Dominican Republic, Ecuador, Egypt, El Salvador, Iraq, Jamaica, Kazakhstan, Sri Lanka, Philippines, Peru, Ukraine
Uruguay	2011	September	Argentina, Belize, Dominican Republic, Ecuador, Egypt, El Salvador, Georgia, Iraq, Jamaica, Kazakhstan, Jordan, Nigeria, Philippines, Peru, Ukraine, Venezuela
Malaysia	2011	September	Argentina, Belize, Dominican Republic, Ecuador, Egypt, El Salvador, Georgia, Iraq, Jamaica, Kazakhstan, Jordan, Nigeria, Philippines, Peru, Ukraine, Venezuela
Mexico	2011	September	Argentina, Belize, Dominican Republic, Ecuador, Egypt, El Salvador, Georgia, Iraq, Jamaica, Kazakhstan, Jordan, Nigeria, Philippines, Peru, Ukraine, Venezuela
Pakistan	2012	January	Argentina, Belize, Dominican Republic, Ecuador, Egypt, El Salvador, Georgia, Iraq, Jamaica, Kazakhstan, Jordan, Nigeria, Philippines, Peru, Ukraine, Venezuela
Turkey	2013	June	Dominican Republic, Ecuador, El Salvador, Iraq, Kazakhstan, Jordan, Nigeria, Pakistan, Sri Lanka, Ukraine, Vietnam
Indonesia	2013	September	Belarus, Belize, Dominican Republic, Ecuador, El Salvador, Iraq, Jordan, Nigeria, Pakistan, Panama, Sri Lanka, Vietnam

Figure 4. Average Treatment Effect Estimates $\hat{\hat{\alpha}}_{1t}$ and Statistical Significance Over Time: Peg-Breaking Episodes Identified with 2.5 Percent Bands



Turkey (2013), and Indonesia (2013). In one case, Chile (2004), the treated country's risk premium decreases more than the synthetic control country's risk premium. In two cases, Mexico (1999) and Malaysia (2011), it does not show either positive or negative treatment effects of breaking pegs.

Figure 4 and Table 7 report the average treatment effect estimates and statistical significance over time.[13] The results mirror those in Section 3.2. The magnitude of the increase in risk premium due to peg abandonment is estimated to be 0.2129–0.3246 standard deviations with a significance level of 99 percent and 0.2010–0.3518

[13]Figure C.1 and Table C.1 in Appendix C present the combined treatment effect estimates using 32 cases (19 cases identified using 2 percent bands and 13 cases (16 cases netting the 3 overlapping cases) identified with 2.5 percent bands). Due to the larger size of treatment effect estimates, the widths of confidence intervals are generally narrower. Therefore, the results indicate that the treatment effect estimates are statistically significant in relatively longer periods after abandoning the exchange rate pegs: four to six months with 95–99 percent significance levels.

Table 7. Average Treatment Effect Estimates $\hat{\hat{\alpha}}_{1t}$ and Statistical Significance Over Time: Peg-Breaking Episodes Identified with 2.5 Percent Bands

Before Switching			After Switching		
t	$\hat{\hat{\alpha}}_{1t}$	p-value	t	$\hat{\hat{\alpha}}_{1t}$	p-value
1	−0.0838	0.3873	T_0	0.2129***	0.0002
2	−0.0123	0.8326	T_1	0.3134***	0.0003
3	0.0456	0.3900	T_2	0.3246***	0.0005
4	0.0024	0.9527	T_3	0.3518**	0.0155
5	−0.0274	0.2625	T_4	0.3104**	0.0266
6	−0.0325	0.3099	T_5	0.2789**	0.0180
7	−0.0299	0.3747	T_6	0.2054**	0.0402
8	0.0090	0.7031	T_7	0.2010*	0.0599
9	0.0635**	0.0464	T_8	0.1523	0.1595
10	0.1066***	0.0054	T_9	0.1417	0.2094
			T	0.1191	0.2585

Note: p-values are calculated for the hypotheses set of $H_0 : \alpha_{1t} = 0$ and $H_1 : \alpha_{1t} \neq 0$. *** indicates 99 percent significance level; ** indicates 95 percent significance level; * indicates 90 percent significance level.

standard deviations with a significance level of 95 percent. The duration of impact is estimated to be from three (99 percent significance level) to seven (95 percent significance level) months.

3.4 Robustness Check: Political Instability

In the previous sections, the peg-breaking episodes that coincide with economic crises were excluded in the estimation of treatment effects because including such cases in the study would have exaggerated the movement of the risk premium index and would result in estimation bias. Political crises may result in similar estimation bias because political instability deters foreign investment in those countries and consequently raises (or amplifies the increase in) the risk premium. If the exchange rate pegs were broken during a period of political turmoil, including such cases in the study potentially results in overestimation of the impact of abandoning the exchange rate peg. In this section, we identify the peg-breaking cases that coincided with political upheavals and provide estimation results excluding those cases.

In order to identify political upheavals, we use the annual political stability index from the Worldwide Governance Indicators (WGI)[14] by the World Bank (2021b). The index measures the perception of the likelihood of political instability in each country. Therefore, the index is well suited to reflect the (local and foreign) investors' sentiment on a country's political stability. As we did with the EMBIG index in (3), we standardize the political stability index to account for country heterogeneity as follows:

$$\mathrm{Std.WGI}_{c,t} = \frac{\mathrm{WGI}_{c,t} - \hat{\mu}(\mathrm{WGI})_c}{\hat{\sigma}(\mathrm{WGI})_c}, \tag{5}$$

where $\mathrm{WGI}_{c,t}$ is the WGI political stability index observation of country c in year t, $\hat{\mu}(\mathrm{WGI})_c$ is the sample average of index of country c, and $\hat{\sigma}(\mathrm{EMBIG})_c$ is the sample standard deviation of index of country c. This standardized index indicates how many standard deviations the country is away from its average level of political stability in year t.

Table 8 presents the standardized political stability index of the year when each country broke their pegs. If the political stability index shows a significant (over one standard deviation[15]) decrease from its average, we assume that the country experienced an extreme political upheaval, which can potentially bias the estimation of risk premium changes during the period; we indicate these cases with the asterisk sign ($*$) in Table 8.

The estimation results excluding these cases show that the main thesis of the study persists. Figures 5 and 6 and Tables 9 and 10 show that the treatment effect remains statistically significant; breaking the exchange rate pegs incurs an increase of 0.2172 to 0.3280

[14] WGI reports six indices to measure aggregate and individual governance for over 200 countries and territories over the period 1996–2019: (i) Voice and Accountability, (ii) Regulatory Quality, (iii) Political Stability and Absence of Violence, (iv) Rule of Law, (v) Government Effectiveness, and (vi) Control of Corruption. In this study, we use (iii) Political Stability and Absence of Violence.

[15] Pakistan's (2012) case is sorted as a political upheaval because the index is almost one standard deviation (0.98 standard deviation) lower than the average.

Table 8. Standardized Political Stability Index in Years of All Cases

Country	Year	Std. WGI
A. Episodes Identified with 2 Percent Bands		
Peru	2001	−0.70
Egypt	2005	0.55
Peru	2005	−0.70
Mexico	2006	−0.21
Ecuador	2010	−0.26
Ghana	2011	1.41
Indonesia	2011	0.49
Pakistan	2012	−0.98*
Egypt	2013	−1.42*
Jamaica	2013	0.76
Argentina	2013	0.53
Indonesia	2013	0.93
Georgia	2014	1.20
Kazakhstan	2014	−0.23
Pakistan	2014	−0.49
Belarus	2015	−0.45
Egypt	2015	−1.14*
Peru	2015	1.06
Philippines	2015	0.85
B. Episodes Identified with 2.5 Percent Bands		
Mexico	1999	0.66
Philippines	2000	−0.40
Korea	2003	−0.88
Argentina	2003	−1.16*
Bulgaria	2003	−0.56
Chile	2004	0.77
Hungary	2005	0.84
Poland	2005	−1.23*
Mexico	2006	−0.21
Belarus	2011	−1.42*
Uruguay	2011	0.67
Malaysia	2011	−0.72
Mexico	2011	−0.33
Pakistan	2012	−0.98*
Turkey	2013	−0.43
Indonesia	2013	0.93

Note: The cases that are considered to coincide with political upheavals are indicated with *.

Figure 5. Average Treatment Effect Estimates $\hat{\hat{\alpha}}_{1t}$ and Statistical Significance Over Time: Excluding Political Crises Cases

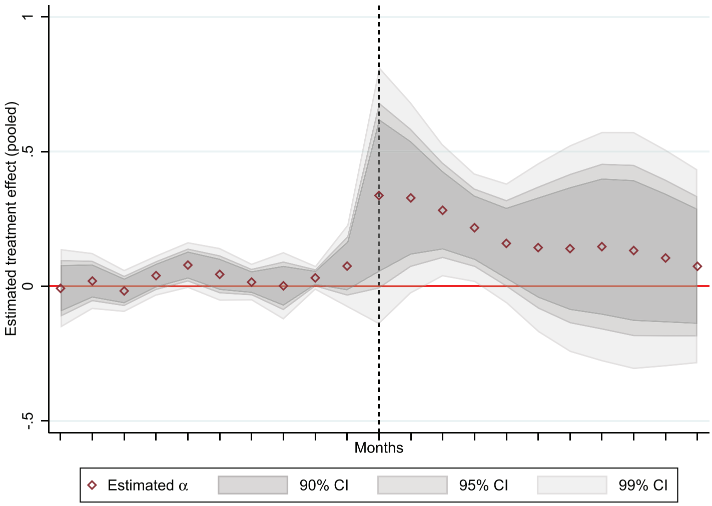


Figure 6. Average Treatment Effect Estimates $\hat{\hat{\alpha}}_{1t}$ and Statistical Significance Over Time: Peg-Breaking Episodes Identified with 2.5 Percent Bands, Excluding Political Crises Cases

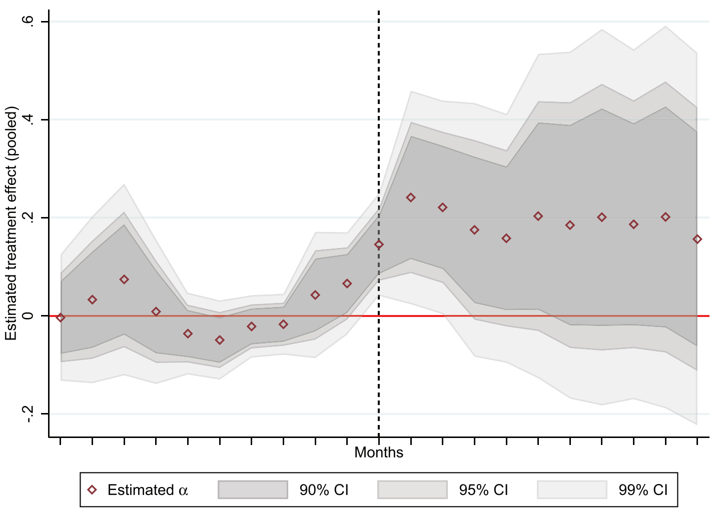


Table 9. Estimated Treatment Effect and Statistical Significance Over Time: Excluding Political Crises Cases

Before Switching			After Switching		
t	$\hat{\hat{\alpha}}_{1t}$	p-value	t	$\hat{\hat{\alpha}}_{1t}$	p-value
1	−0.0079	0.8760	T_0	0.3369*	0.0558
2	0.0193	0.5954	T_1	0.3280**	0.0159
3	−0.0175	0.5226	T_2	0.2820***	0.0042
4	0.0394	0.1437	T_3	0.2172***	0.0064
5	0.0786**	0.0164	T_4	0.1591*	0.0531
6	0.0441	0.2075	T_5	0.1434	0.1988
7	0.0153	0.5235	T_6	0.1397	0.3011
8	0.0014	0.9735	T_7	0.1470	0.3259
9	0.0310*	0.0628	T_8	0.1326	0.3894
10	0.0748	0.1701	T_9	0.1049	0.4553
			T	0.0739	0.5550

Note: *** indicates 99 percent significance level; ** indicates 95 percent significance level; * indicates 90 percent significance level.

Table 10. Estimated Treatment Effect and Statistical Significance Over Time: Peg-Breaking Episodes Identified with 2.5 Percent Bands, Excluding Political Crises Cases

Before Switching			After Switching		
t	$\hat{\hat{\alpha}}_{1t}$	p-value	t	$\hat{\hat{\alpha}}_{1t}$	p-value
1	−0.0035	0.9339	T_0	0.1456***	0.0013
2	0.0330	0.5606	T_1	0.2414***	0.0055
3	0.0743	0.2635	T_2	0.2212***	0.0092
4	0.0086	0.8595	T_3	0.1753*	0.0594
5	−0.0363	0.2048	T_4	0.1581*	0.0794
6	−0.0493*	0.0850	T_5	0.2034*	0.0827
7	−0.0216	0.3158	T_6	0.1850	0.1329
8	−0.0172	0.4109	T_7	0.2012	0.1321
9	0.0426	0.3278	T_8	0.1867	0.1326
10	0.0659*	0.0760	T_9	0.2017	0.1370
			T	0.1564	0.2274

Note: *** indicates 99 percent significance level; ** indicates 95 percent significance level; * indicates 90 percent significance level.

standard deviations[16] in the country risk spread (95–99 percent significance levels), and the estimated treatment effects are statistically positive for two to three months.[17] Although the duration of treatment effect may seem to decrease, the decrease is potentially due to the widened confidence intervals from the decrease in number of samples by excluding three to four cases; Figures 5 and 6 show that the estimated treatment effects still "jump" and stay positive after the exchange rate pegs are broken, as we have seen in Figure 3.

3.5 Robustness Check: Peg Formation

The study hypothesizes that investors may interpret breaking exchange rate pegs as the central banks' loss in the ability to keep their commitment to monetary stability. In order to bolster the hypothesis, this section investigates episodes of exchange rate peg formation. Based on the hypothesis, switching regime from fixed to floating and floating to fixed should have an asymmetric impact on risk premium. Investor confidence may falter, resulting in a decrease in investment in the country immediately after a breaking of the peg, which can be interpreted as the central bank breaking a commitment. On the other hand, central banks gaining investors' trust after the formation of a peg may take some time. This exercise will therefore confirm the hypothesis if the treatment effect (of switching the regime from floating to peg) does not drop below zero or show any systemic changes after peg formation.

We use the same basis of coding used in Section 2.2 to identify de facto regime periods; we classify a country as intervening in the foreign exchange market if the exchange rates stay within a 2 percent band for 10 consecutive months. We isolate episodes where a country floats its exchange rates for at least 10 months before the de facto fixed regime begins. The control country candidates are the countries which keep de facto floating regimes during the study period. We identify 33 peg-formation episodes, listed with the control country candidates for each episode in Table 11.

[16] These results are derived from using the events and countries identified with 2 percent bands; 0.1456–0.2414 standard deviation increases are estimated using 2.5 percent bands.

[17] Similar results can be found in combined placebo tests, as shown in Figure E.2 in Appendix E.

Table 11. Thirty-Three Cases of Peg Formation and Control Country Candidates to Construct Synthetic Controls

Country	Year	Month	Control Country Candidates
Korea	2000	January	Algeria, Brazil, Bulgaria, Colombia, Côte d'Ivoire, Croatia, Hungary, Mexico, Morocco, Nigeria, Poland, Thailand, Turkey
Ecuador	2000	February	Algeria, Brazil, Bulgaria, Colombia, Côte d'Ivoire, Croatia, Hungary, Mexico, Morocco, Nigeria, Poland, Thailand, Turkey
Colombia	2001	February	Brazil, Bulgaria, Chile, Côte d'Ivoire, Croatia, Hungary, Mexico, Morocco, Nigeria, Philippines, Poland, Turkey
Thailand	2001	May	Brazil, Bulgaria, Chile, Côte d'Ivoire, Croatia, Hungary, Mexico, Morocco, Nigeria, Philippines, Poland, Turkey
Algeria	2001	September	Brazil, Bulgaria, Chile, Côte d'Ivoire, Croatia, Hungary, Mexico, Morocco, Nigeria, Philippines, Poland, Turkey
Venezuela	2003	April	Argentina, Brazil, Chile, Côte d'Ivoire, Croatia, Dominican Republic, Hungary, Morocco, Poland, Turkey, Uruguay
Colombia	2003	June	Argentina, Brazil, Chile, Côte d'Ivoire, Croatia, Dominican Republic, Hungary, Morocco, Poland, Turkey, Uruguay
Philippines	2003	September	Brazil, Bulgaria, Chile, Côte d'Ivoire, Dominican Republic, Hungary, Morocco, Poland, Turkey, Uruguay
Nigeria	2003	December	Brazil, Bulgaria, Chile, Côte d'Ivoire, Dominican Republic, Hungary, Morocco, Poland, Turkey, Uruguay
Argentina	2004	June	Brazil, Bulgaria, Chile, Côte d'Ivoire, Dominican Republic, Hungary, Morocco, Poland, Turkey, Uruguay
Mexico	2004	June	Brazil, Bulgaria, Chile, Côte d'Ivoire, Dominican Republic, Hungary, Morocco, Poland, Turkey, Uruguay
Turkey	2005	May	Brazil, Bulgaria, Chile, Côte d'Ivoire, Dominican Republic, Hungary, Indonesia, Poland
Morocco	2005	July	Brazil, Bulgaria, Chile, Côte d'Ivoire, Hungary, Indonesia, Poland
Uruguay	2006	February	Brazil, Bulgaria, Chile, Côte d'Ivoire, Hungary, Indonesia, Poland
Dominican Rep.	2006	April	Brazil, Bulgaria, Chile, Côte d'Ivoire, Hungary, Indonesia, Poland
Ghana	2009	June	Brazil, Bulgaria, Colombia, Chile, Gabon, Hungary, Malaysia, Mexico, Philippines, Poland, Turkey, Uruguay

(*continued*)

Table 11. (Continued)

Country	Year	Month	Control Country Candidates
Peru	2009	June	Brazil, Bulgaria, Colombia, Chile, Gabon, Hungary, Malaysia, Mexico, Philippines, Poland, Turkey, Uruguay
Indonesia	2009	November	Brazil, Bulgaria, Colombia, Chile, Gabon, Hungary, Malaysia, Mexico, Philippines, Poland, Turkey
Ukraine	2009	November	Brazil, Bulgaria, Colombia, Chile, Gabon, Hungary, Malaysia, Mexico, Philippines, Poland, Turkey
Uruguay	2010	August	Brazil, Bulgaria, Colombia, Chile, Gabon, Hungary, Malaysia, Mexico, Poland, Turkey
Philippines	2010	October	Brazil, Bulgaria, Colombia, Chile, Croatia, Gabon, Hungary, Malaysia, Mexico, Poland, Turkey
Georgia	2011	May	Brazil, Bulgaria, Chile, Côte d'Ivoire, Croatia, Gabon, Hungary, Malaysia, Mexico, Poland
Colombia	2012	March	Brazil, Bulgaria, Chile, Côte d'Ivoire, Croatia, Gabon, Hungary, Mexico, Poland, Senegal
Turkey	2012	March	Brazil, Bulgaria, Chile, Côte d'Ivoire, Croatia, Gabon, Hungary, Mexico, Poland, Senegal
Ghana	2012	June	Brazil, Bulgaria, Chile, Côte d'Ivoire, Croatia, Gabon, Hungary, Mexico, Poland, Senegal
Malaysia	2012	July	Brazil, Bulgaria, Chile, Côte d'Ivoire, Croatia, Gabon, Hungary, Mexico, Poland, Senegal
Croatia	2013	July	Brazil, Chile, Hungary, Poland, Uruguay
Mexico	2013	July	Brazil, Chile, Hungary, Poland, Uruguay
Côte d'Ivoire	2013	November	Brazil, Chile, Hungary, Poland, Uruguay
Gabon	2013	November	Brazil, Chile, Hungary, Poland, Uruguay
Senegal	2013	November	Brazil, Chile, Hungary, Poland, Uruguay
Jamaica	2014	February	Brazil, Chile, Hungary, Poland, Uruguay
Nigeria	2015	March	Brazil, Colombia, Chile, Ghana, Hungary, Indonesia, Poland, Turkey, Ukraine, Uruguay

Figure 7 and Table 12 show the average treatment effect estimates of peg formation and their statistical significance over time.[18] The estimation results confirm the hypothesis that the "benefit" of peg formation and the "cost" of breaking a peg are asymmetric. That is, while the previous results confirm that positive and significant costs of breaking an exchange rate peg exist, the results

[18]The individual synthetic control analysis results and placebo tests are presented in Figures D.1 and D.2 in Appendix D.

Figure 7. Estimated Treatment Effect of Foreign Exchange Regime Switching on Country's Risk Premium Over Time: Peg-Formation Episodes

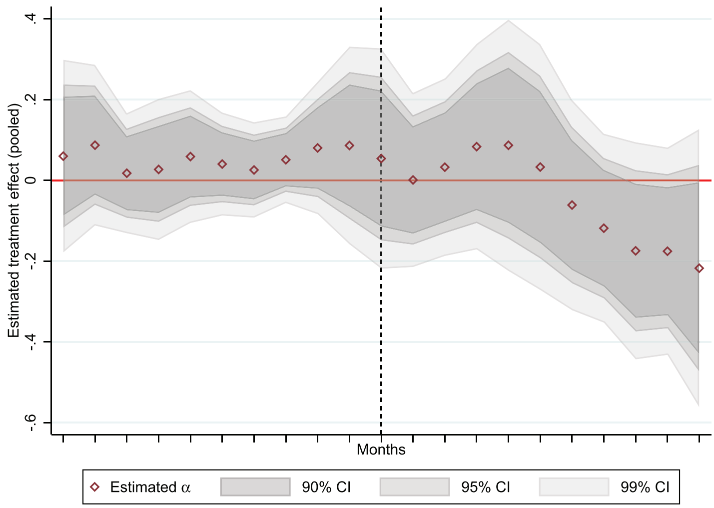


Table 12. Estimated Treatment Effect of Peg Formation and Statistical Significance Over Time

Before Switching			After Switching		
t	$\hat{\hat{\alpha}}_{1t}$	p-value	t	$\hat{\hat{\alpha}}_{1t}$	p-value
1	0.0599	0.4973	T_0	0.0539	0.5924
2	0.0870	0.2405	T_1	0.0009	0.9905
3	0.0177	0.7475	T_2	0.0326	0.6880
4	0.0270	0.6756	T_3	0.0834	0.3766
5	0.0588	0.3357	T_4	0.0868	0.4511
6	0.0403	0.3941	T_5	0.0329	0.7693
7	0.0257	0.5568	T_6	−0.0612	0.5251
8	0.0509	0.2038	T_7	−0.1185	0.1752
9	0.0802	0.1905	T_8	−0.1745*	0.0851
10	0.0863	0.3419	T_9	−0.1756*	0.0704
			T	−0.2177*	0.0936

Note: *** indicates 99 percent significance level; ** indicates 95 percent significance level; * indicates 90 percent significance level.

in this section suggest that peg formation does not have a statistical impact on sovereign risk spreads. The average treatment effect estimates do not deviate much from zero, and they stay statistically insignificant for the entire study period with 95–99 percent significance levels (for eight months with 90 percent level). It is also noteworthy that the downward trend seems to begin at the end of the study period; that is, stable exchange rates do not lower the risk premium unless they last longer than nine months (with 90 percent significance level). In other words, the results imply that while breaking a peg is interpreted as breaking the central bank's commitment, which is reflected as an immediate increase (or "jump") in the country risk spread, it may take time to gain investors' trust in the central bank's commitment to stabilize the market and the economy.

3.6 Further Discussion of the Results

Although the results in the study confirm the existence of statistically significant short-term costs of breaking a peg, the study also suggests that the effect is potentially short-lived. After its spike from the exchange rate regime switch, the statistical significance of the average treatment effect tends to dissipate within a few months. The estimated short-term effect has two potential explanations. One explanation is that the average treatment effect decreases over time, as seen in Figure 3 and Table 5. However, they also show that the average treatment effect is still above zero for the entire study period of 10 months. Another potential explanation is that the variance in the estimated risk premium increases after the peg breaking. In other words, the change in the exchange rate regime may affect different countries differently, and the magnitude and the duration of impacts differ by country. The difference in impacts creates rather large standard errors and thus wide confidence intervals. Therefore, even if the average treatment effect remains positive, the estimated treatment effect becomes statistically zero after four to seven months, due to the large standard errors.

While this study focuses on the identification of the costs of breaking a peg in the short run, a further investigation of the treatment effect in the long run can be beneficial. A potential extension of the study is to examine the determinants of the impact of peg breaking on risk premium. The treatment effects vary among peg-breaking episodes, and such variation in the magnitude and the duration of impacts exists even within a country between different peg-breaking episodes; for example, while the risk premium increase is estimated to be around one standard deviation in the case of Peru in 2001, a much smaller increase (0.1 standard deviation) is estimated when the peg was broken in 2015. Identifying the determinants of such variation in treatment effects potentially leads us to understand why it may cost more for specific countries (or at certain times) to abandon their exchange rate pegs. The results of research projects such as these would contribute to our understanding of the rationality for central banks to fear to float in the long run. We leave these extensions for future work.

4. Conclusion

This paper studies the short-term cost of switching the exchange rate regime from fixed to floating by investigating the country spread through the synthetic control method. We use J.P. Morgan's risk premium index to measure the country spread that is independent of exchange risks to capture the foreign investor's sentiment on the sovereign risk during exchange rate peg abandonment. Overall, the results provide clear empirical evidence that switching exchange rate regimes from fixed to floating incurs an increase in the risk premium. When the exchange rate peg is broken, the average risk premium shows a statistically significant increase of 0.2129–0.3246 standard deviations. The results support the hypothesis that breaking exchange rate pegs may bring negative sentiment to the market, increasing the price of foreign loans, which is reflected in the statistically significant increase in sovereign risk.

In order to check the robustness of the results, we repeat the analysis on peg-breaking episodes identified with an alternative specification used to classify exchange rate regimes; the treatment effect is estimated to be an increase of 0.2010–0.3518 standard deviations in the country risk spread for three to seven months. In addition, we present the estimation results excluding the cases that coincided with political crises, which may incur overestimation of treatment effects, and the estimation without these cases provides similar results: an increase of 0.2172–0.3280 standard deviations in the risk premium. In both robustness check exercises, statistically significant country risk premium increases are estimated after countries abandon their exchange rate pegs. We also investigate episodes of peg formation. The results confirm the hypothesis of Alesina and Wagner (2006) and suggest that breaking a peg implies breaking the central bank's commitment and that it may take a long time for the central bank to regain investors' trust in its commitment to monetary stability.

Appendix A. List of Peg-Breaking Episodes that Coincide with Economic Crises

Table A.1. Peg-Breaking Cases that are Excluded in Estimation Due to the Coincidence with Economic Crises

Exchange Rate Peg-Breaking Cases			
Country	**Year**	**Month**	**Economic Crises**
Argentina	2002	February	Argentine Great Depression (1998–2002)
Venezuela	2002	March	Venezuelan General Strike (2002–03)
Iraq	2007	January	Subprime Mortgage Crisis (2007–09)
Malaysia	2007	January	
Philippines	2007	June	
Uruguay	2007	November	
Peru	2007	November	
Pakistan	2008	June	
Argentina	2008	July	
Jamaica	2008	November	
Venezuela	2010	February	Venezuelan Banking Crisis (2009–10)
Venezuela	2013	March	Crisis in Venezuela (2012–)
Ukraine	2014	March	Ukrainian Crisis (2013–14)

Appendix B. Synthetic Control Analyses on Peg-Breaking Episodes Identified with 2.5 Percent Bands de facto Regime Specification

Table B.1. Predictor Balance in Synthetic Country Construction: Peg-Breaking Episodes Identified with 2.5% Bands

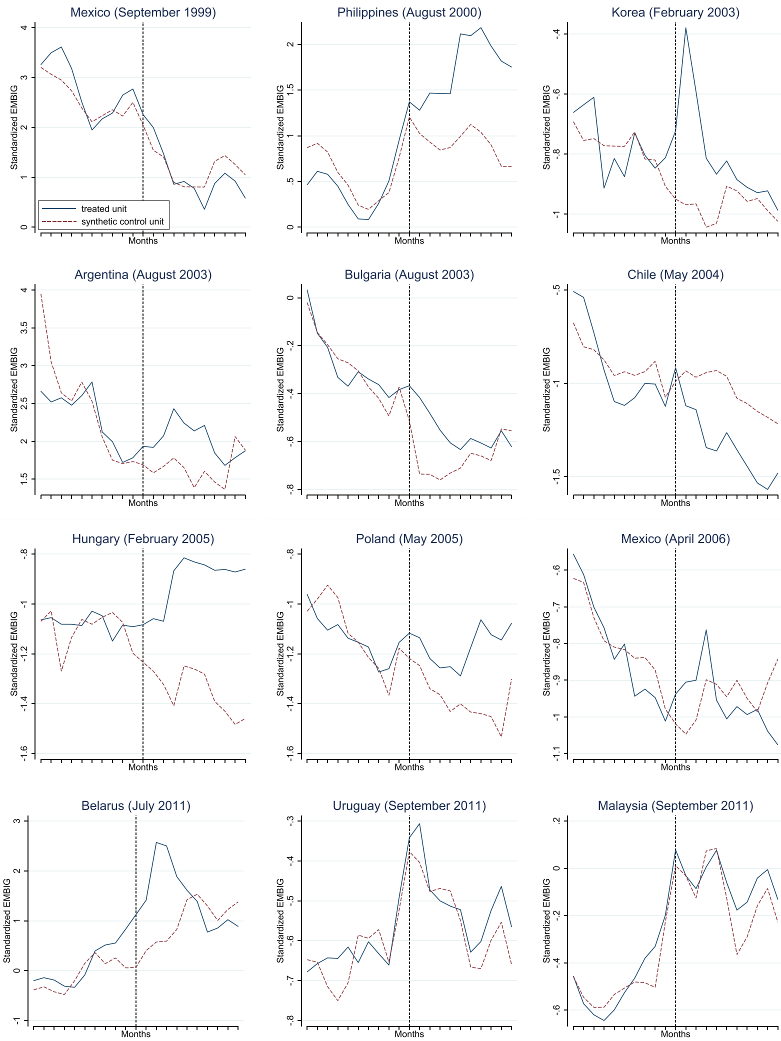
	Treated	Synthetic		Treated	Synthetic		Treated	Synthetic
Mexico, Sept. 1999			**Hungary, Feb. 2005**			**Mexico, Sept. 2011**		
Consumer Price Index	4.072	4.413	Consumer Price Index	4.331	4.174	Consumer Price Index	4.648	4.640
Imports Merchandise	9.690	7.473	Imports Merchandise	8.657	7.464	Imports Merchandise	10.184	7.883
Exports Merchandise	9.332	6.915	Exports Merchandise	8.462	6.755	Exports Merchandise	10.260	8.265
Months Import Cover	–1.284	1.703	Months Import Cover	0.978	1.789	Months Import Cover	–1.022	2.723
Total Reserves	8.085	8.854	Industrial Production	21.635	22.001	Industrial Production	24.126	22.349
RMSPE		0.757	Total Reserves	9.526	9.047	Total Reserves	9.244	10.722
			RMSPE		0.177	**RMSPE**		0.085
Philippines, Aug. 2000			**Poland, Feb. 2005**			**Pakistan, Jan. 2012**		
Imports Merchandise	7.972	7.395	Consumer Price Index	4.469	4.086	Consumer Price Index	4.769	4.769
Exports Merchandise	8.030	7.217	Imports Merchandise	9.251	7.448	Imports Merchandise	8.061	8.057
Months Import Cover	1.712	2.344	Exports Merchandise	8.846	6.617	Exports Merchandise	7.672	7.438
Industrial Production	22.552	22.443	Months Import Cover	1.555	2.401	Months Import Cover	1.627	1.627
Total Reserves	9.626	9.482	Industrial Production	22.899	22.306	Total Reserves	9.811	9.811
RMSPE		0.256	Total Reserves	10.548	9.646	**RMSPE**		0.098
			RMSPE		0.083			
Korea, Feb. 2003			**Mexico, Apr. 2006**			**Turkey, Jun. 2013**		
Consumer Price Index	4.377	4.242	Consumer Price Index	4.415	4.259	Consumer Price Index	4.816	4.816
Imports Merchandise	10.049	8.424	Imports Merchandise	10.009	7.479	Imports Merchandise	9.925	8.403
Exports Merchandise	9.549	8.077	Exports Merchandise	9.842	7.073	Exports Merchandise	9.457	8.943
Months Import Cover	2.165	1.262	Months Import Cover	1.165	1.725	Months Import Cover	1.778	2.529
Total Reserves	11.659	9.417	Industrial Production	24.113	21.951	Total Reserves	11.704	10.933
RMSPE		0.087	Total Reserves	8.695	9.054	**RMSPE**		0.654
			RMSPE		0.031			

(*continued*)

Table B.1. (Continued)

	Treated	Synthetic		Treated	Synthetic		Treated	Synthetic
Argentina, Aug. 2003			**Belarus, Jul. 2011**			**Indonesia, Sept. 2013**		
Imports Merchandise	7.175	7.207	Consumer Price Index	4.768	4.722	Consumer Price Index	4.763	4.763
Exports Merchandise	7.797	6.907	Imports Merchandise	8.190	8.428	Imports Merchandise	9.669	8.354
Months Import Cover	2.424	2.242	Exports Merchandise	7.944	7.627	Exports Merchandise	9.628	8.777
Total Reserves	9.287	9.155	Months Import Cover	–0.138	1.940	Months Import Cover	1.891	2.457
RMSPE		1.027	Industrial Production	21.651	22.651	Total Reserves	11.559	10.811
			Total Reserves	8.052	10.368	**RMSPE**		0.403
			RMSPE		0.325			
Bulgaria, Aug. 2003			**Uruguay, Sept. 2011**					
Imports Merchandise	6.760	7.818	Consumer Price Index	4.694	4.639			
Exports Merchandise	6.453	5.789	Imports Merchandise	6.730	7.763			
Months Import Cover	1.890	1.897	Exports Merchandise	–0.450	8.012			
Industrial Production	20.607	21.902	Months Import Cover	9.146	2.101			
Total Reserves	8.659	9.715	Industrial Production	20.608	22.091			
RMSPE		0.054	Total Reserves	9.052	9.973			
			RMSPE		0.058			
Chile, May 2004			**Malaysia, Sept. 2011**					
Imports Merchandise	7.623	7.625	Consumer Price Index	4.635	4.671			
Exports Merchandise	7.665	7.004	Imports Merchandise	9.494	7.920			
Months Import Cover	2.754	2.330	Exports Merchandise	9.844	8.385			
Industrial Production	22.475	22.160	Months Import Cover	2.054	2.472			
Total Reserves	10.127	9.702	Industrial Production	22.879	22.391			
RMSPE		0.140	Total Reserves	11.700	10.529			
			RMSPE		0.076			

Figure B.1. Trends of Standardized EMBIG of Treated Country (solid line) vs. Synthetic Control (dashed line): Peg-Breaking Episodes Identified with 2.5 Percent Bands



(continued)

Figure B.1. (Continued)

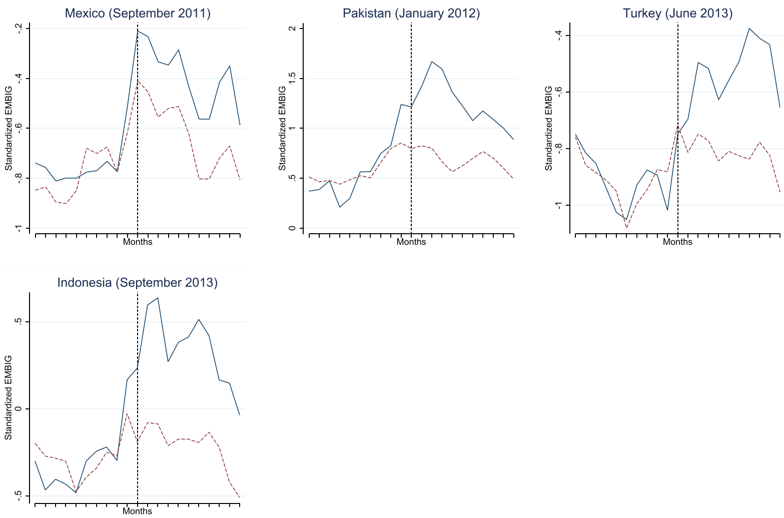
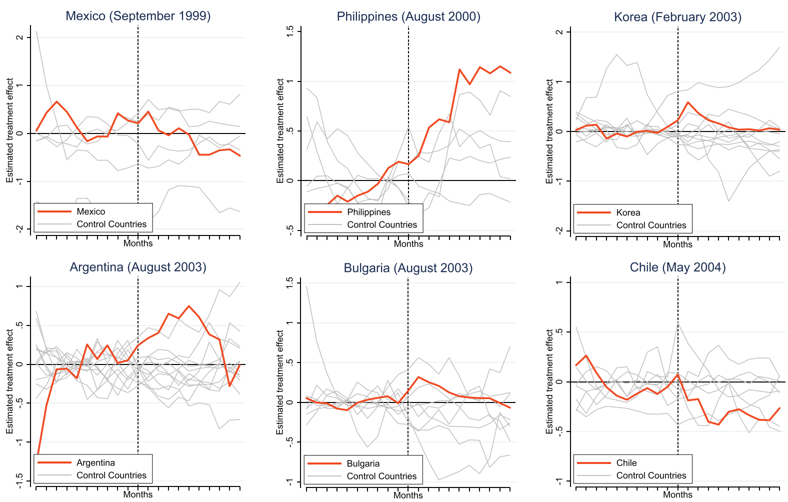
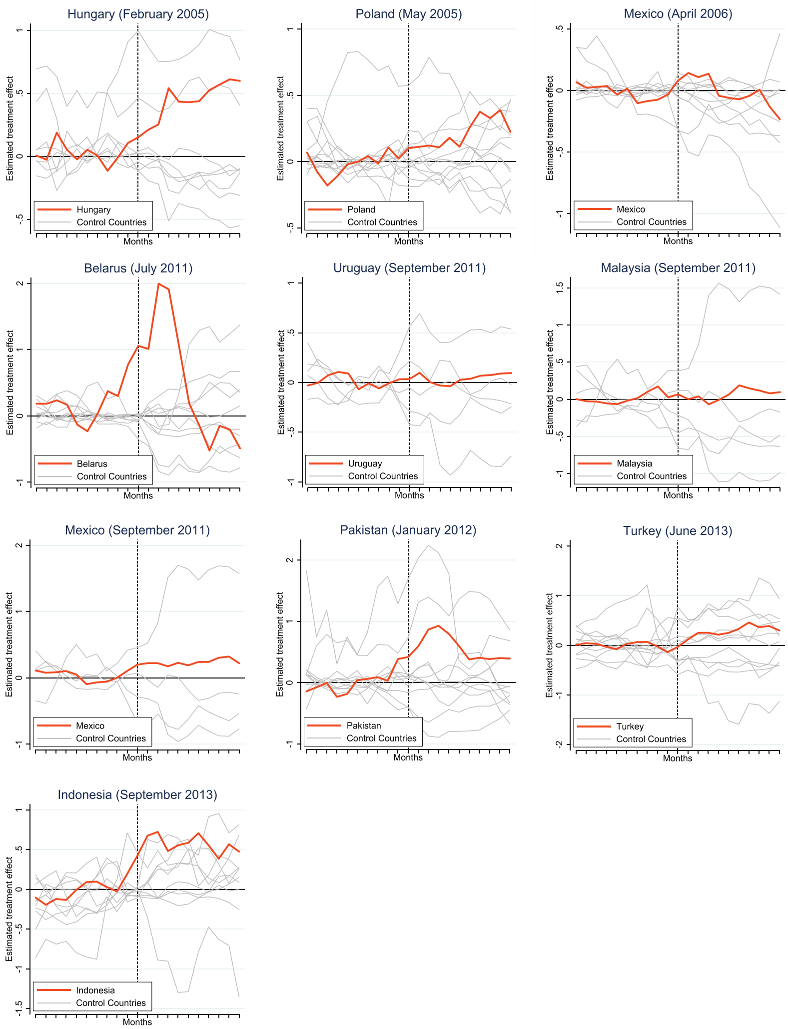


Figure B.2. Trends of Estimated Treatment Effect (orange line) vs. Placebo Effect (gray line): Peg-Breaking Episodes Identified with 2.5 Percent Bands



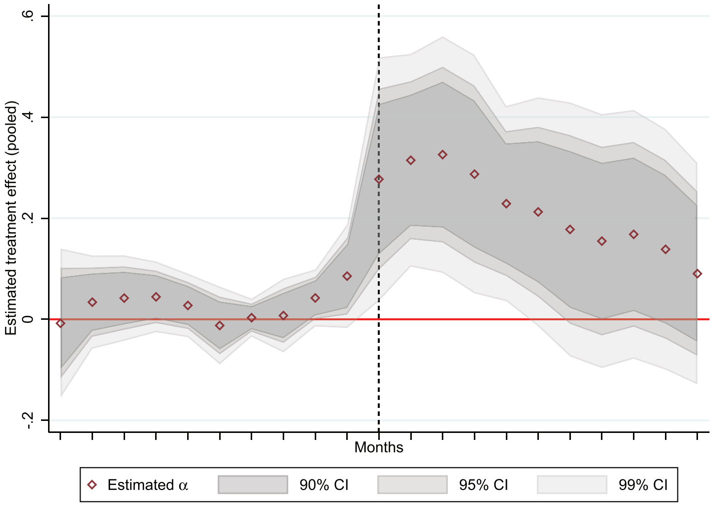
(continued)

Figure B.2. (Continued)



Appendix C. The Estimated Treatment Effect: Using 32 Peg-Breaking Cases Identified with Both 2 Percent and 2.5 Percent Bands

Figure C.1. Estimated Treatment Effect and Statistical Significance Over Time



Note: The figure shows 32 episodes identified with 2 percent (19 cases) and 2.5 percent bands (13 cases, excluding the overlapping cases of Mexico (2006), Pakistan (2013), and Indonesia (2013)).

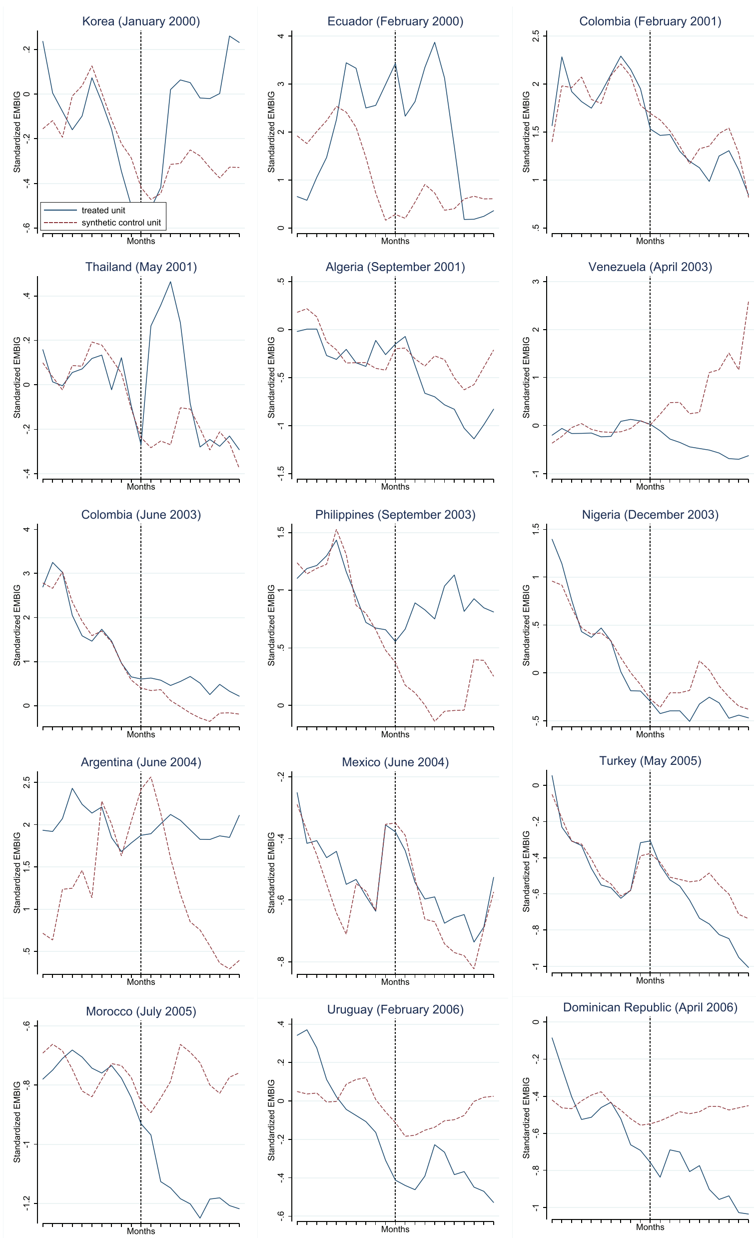
Table C.1. Estimated Treatment Effect and Statistical Significance Over Time: Using the 32 Episodes Identified with 2 Percent (19 cases) and 2.5 Percent (13 cases: excluding the overlapping cases of Mexico (2006), Pakistan (2013), and Indonesia (2013))

Before Switching			After Switching		
t	$\hat{\hat{\alpha}}_{1t}$	p-value	t	$\hat{\hat{\alpha}}_{1t}$	p-value
1	−0.0079	0.8853	T_0	0.2773**	0.0036
2	0.0339621	0.3215	T_1	0.3148***	0.0003
3	0.0419931	0.1837	T_2	0.3259***	0.0006
4	0.0444137	0.0923	T_3	0.2873***	0.0022
5	0.0274539	0.2413	T_4	0.2289***	0.0028
6	−0.0122	0.6686	T_5	0.2126**	0.0152
7	0.0032	0.8170	T_6	0.1778*	0.0615
8	0.0074	0.7833	T_7	0.1548	0.1012
9	0.0423**	0.0493	T_8	0.1682	0.0704
10	0.0855**	0.0294	T_9	0.1385	0.1211
			T	0.0904	0.2670

Note: *** indicates 99 percent significance level; ** indicates 95 percent significance level; * indicates 90 percent significance level.

Appendix D. Synthetic Control Analyses on the Episodes of Peg Formation

Figure D.1. Trends of Standardized EMBIG of Treated Country (solid line) vs. Synthetic Control (dashed line): Peg-Formation Episodes



(continued)

Figure D.1. (Continued)

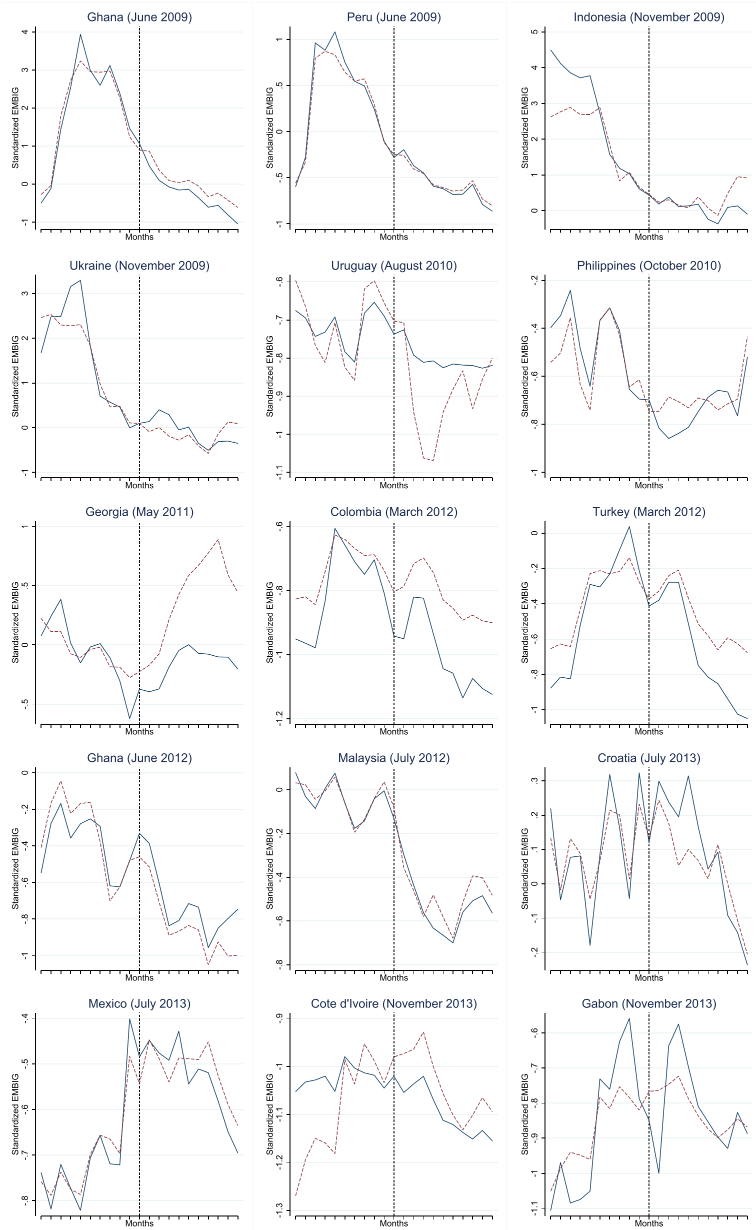
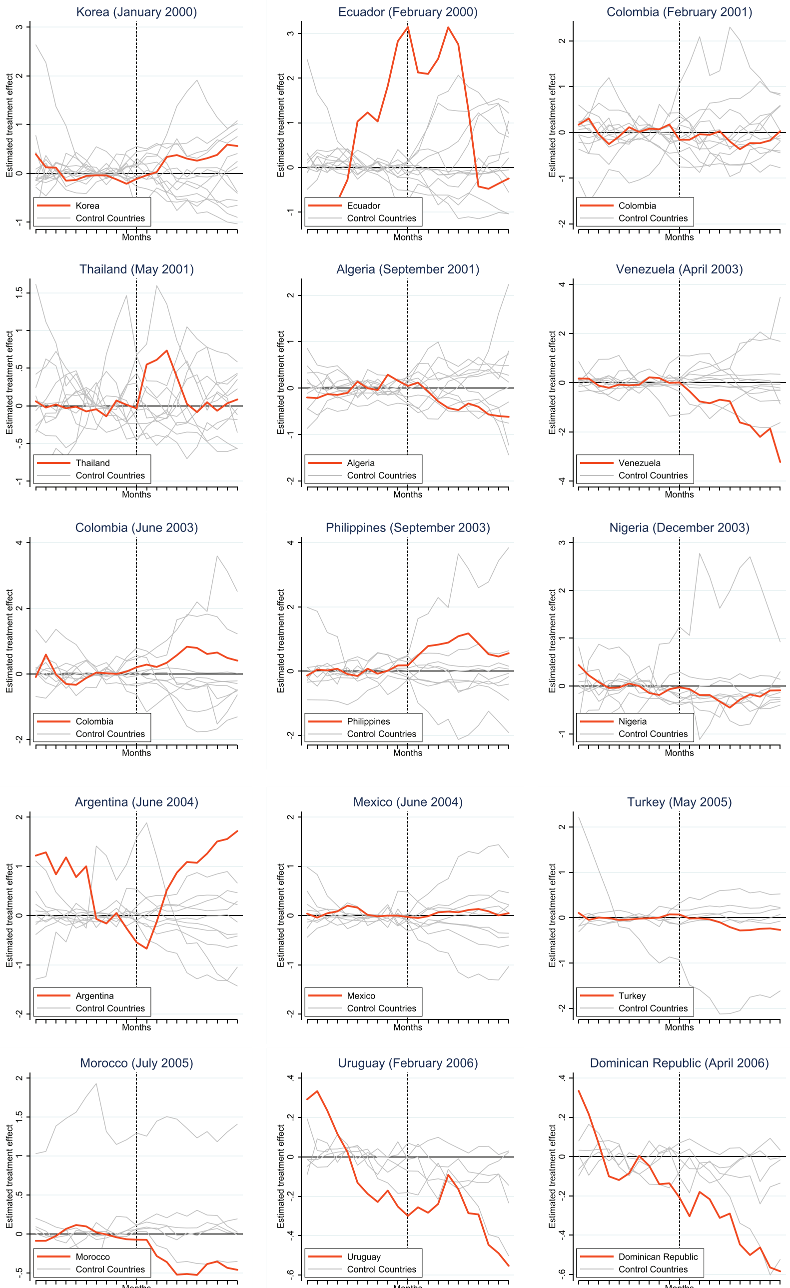
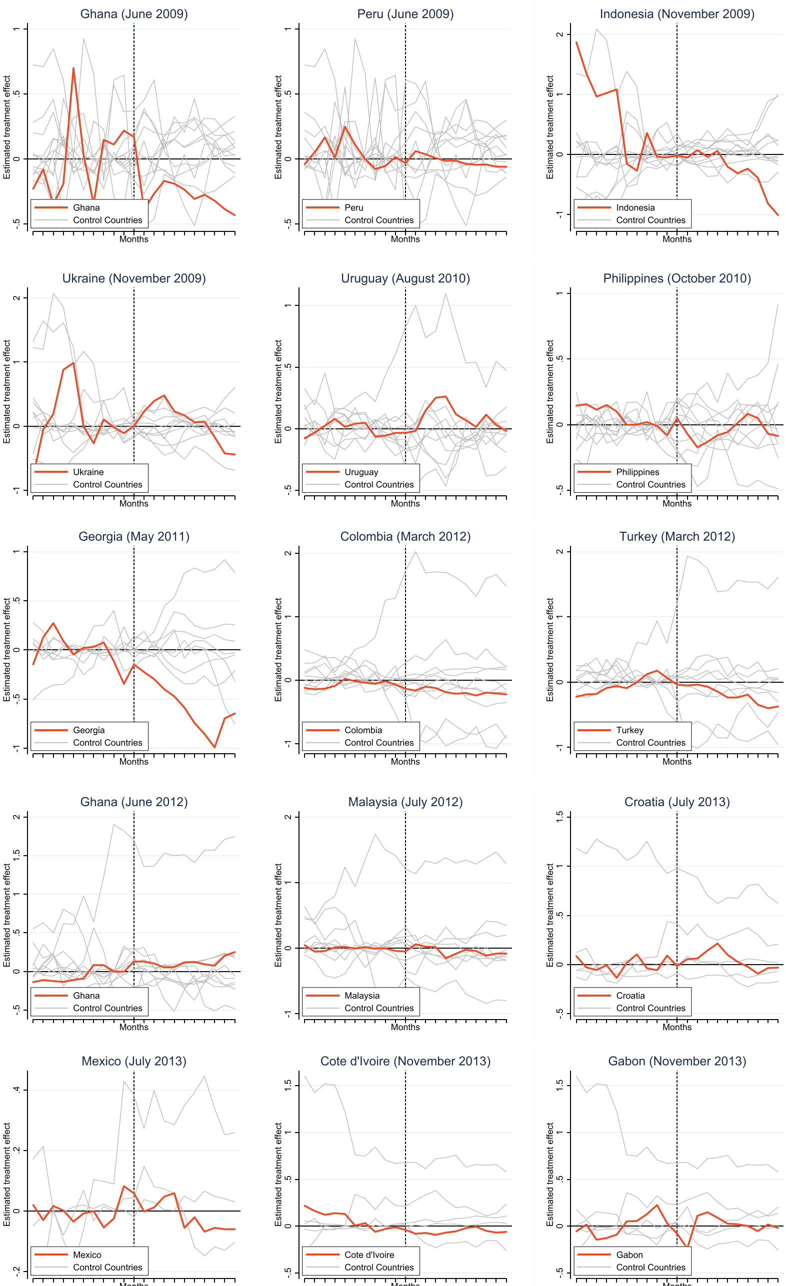


Figure D.2. Trends of Estimated Treatment Effect (orange line) vs. Placebo Effect (gray line): Peg-Formation Episodes



(continued)

Figure D.2. (Continued)



Appendix E. Combined Placebo Tests

Figure E.1. Combined Placebo Tests on Control Countries

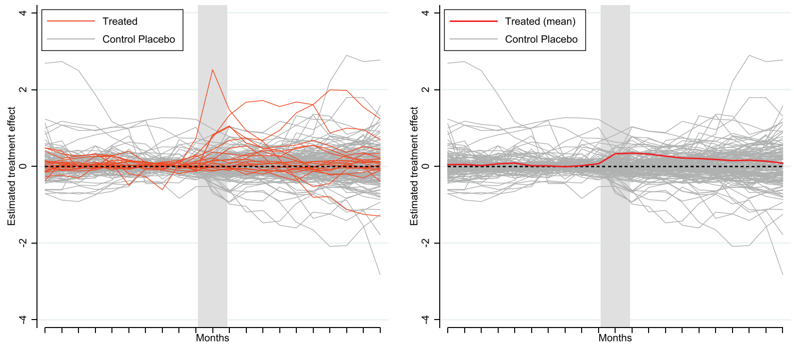
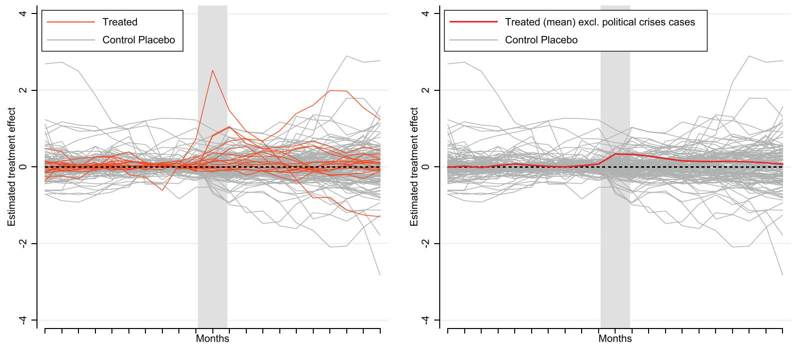


Figure E.2. Combined Placebo Tests (excluding the results of political crises cases) on Control Countries



References

Abadie, A., A. Diamond, and J. Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105 (490): 493–505.

———. 2015. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science* 59 (2): 495–510.

Abadie, A., and J. Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review* 93 (1): 113–32.

Aghion, P., P. Bacchetta, R. Ranciere, and K. Rogoff. 2009. "Exchange Rate Volatility and Productivity Growth: The Role of Financial Development." *Journal of Monetary Economics* 56 (4): 494–513.

Aizenman, J. 2013. "The Impossible Trinity—From the Policy Trilemma to the Policy Quadrilemma." *Global Journal of Economics* 2 (1): Article 1350001.

Alesina, A., and A. F. Wagner. 2006. "Choosing (and Reneging on) Exchange Rate Regimes." *Journal of the European Economic Association* 4 (4): 770–99.

Batini, N., P. Breuer, K. Kochhar, and S. Roger. 2006. "Inflation Targeting and the IMF." International Monetary Fund.

Billmeier, A., and T. Nannicini. 2013. "Assessing Economic Liberalization Episodes: A Synthetic Control Approach." *Review of Economics and Statistics* 95 (3): 983–1001.

Calvo, G. A., and C. M. Reinhart. 2002. "Fear of Floating." *Quarterly Journal of Economics* 117 (2): 379–408.

Ferman, B., and C. Pinto. 2017. "Placebo Tests for Synthetic Controls." Technical report, University Library of Munich, Germany.

Gapen, M., D. Gray, C. H. Lim, and Y. Xiao. 2008. "Measuring and Analyzing Sovereign Risk with Contingent Claims." *IMF Staff Papers* 55 (1): 109–48.

Ghosh, A. R., A. M. Gulde, and H. C. Wolf. 2002. *Exchange Rate Regimes: Choices and Consequences*, Vol. 1. MIT Press.

Hahn, J., and R. Shi. 2017. "Synthetic Control and Inference." *Econometrics* 5 (4): 52.

Harms, P., and M. Kretschmann. 2009. "Words, Deeds and Outcomes: A Survey on the Growth Effects of Exchange Rate Regimes." *Journal of Economic Surveys* 23 (1): 139–64.

Hilscher, J., and Y. Nosbusch. 2010. "Determinants of Sovereign Risk: Macroeconomic Fundamentals and the Pricing of Sovereign Debt." *Review of Finance* 14 (2): 235–62.

Jahjah, S., B. Wei, and V. Z. Yue. 2013. "Exchange Rate Policy and Sovereign Bond Spreads in Developing Countries." *Journal of Money, Credit and Banking* 45 (7): 1275–1300.

Kaminsky, G., and S. L. Schmukler. 2002. "Emerging Market Instability: Do Sovereign Ratings Affect Country Risk and Stock Returns?" *World Bank Economic Review* 16 (2): 171–95.

Klein, M. W., and J. C. Shambaugh. 2008. "The Dynamics of Exchange Rate Regimes: Fixes, Floats, and Flips." *Journal of International Economics* 75 (1): 70–92.

Levy-Yeyati, E., and F. Sturzenegger. 2003. "To Float or to Fix: Evidence on the Impact of Exchange Rate Regimes on Growth." *American Economic Review* 93 (4): 1173–93.

Mishkin, F. S. 1999. "Lessons from the Asian Crisis." *Journal of International Money and Finance* 18 (4): 709–23.

Mundell, R. A. 1963. "Capital Mobility and Stabilization Policy under Fixed and Flexible Exchange Rates." *Canadian Journal of Economics and Political Science/Revue canadienne de economiques et science politique* 29 (4): 475–85.

Popper, H., A. Mandilaras, and G. Bird. 2013. "Trilemma Stability and International Macroeconomic Archetypes." *European Economic Review* 64 (November): 181–93.

Remolona, E. M., M. Scatigna, and E. Wu. 2008. "The Dynamic Pricing of Sovereign Risk in Emerging Markets: Fundamentals and Risk Aversion." *Journal of Fixed Income* 17 (4): 57–71.

World Bank. 2021a. "Global Economic Monitor." Data retrieved from https://datacatalog.worldbank.org/dataset/global-economic-monitor.

World Bank. 2021b. "The Worldwide Governance Indicators." Data retrieved from the Worldwide Governance Indicators, https://info.worldbank.org/governance/wgi/.

Towards a Macroprudential Framework for Investment Funds: Swing Pricing and Investor Redemptions*

Ulf Lewrick[a,b] and Jochen Schanz[c]
[a]Bank for International Settlements
[b]University of Basel
[c]European Investment Bank

How effective are available policy tools in managing systemic liquidity risks in the mutual fund industry? We assess one such tool—swing pricing—which allows funds to adjust their settlement price in response to large flows. A global game guides our empirical analysis. Consistent with its predictions, we show that during normal market conditions swing pricing dampens outflows in reaction to weak fund performance by mitigating investor first-mover advantages. Yet during episodes of market stress, swing pricing fails to contain redemption pressures despite supporting fund returns. This calls for adjusting swing pricing rules to achieve macroprudential objectives.

JEL Codes: G01, G23, G28, C72.

*The views expressed in this article are those of the authors and do not necessarily reflect those of the Bank for International Settlements or of the European Investment Bank. For their helpful comments, we thank Elena Loutskina (the editor), an anonymous referee, Sirio Aramonte, Claudio Borio, Stijn Claessens, Andrew Ellul, Ingo Fender, Itay Goldstein, Leslie Kapin, Gianpaolo Parise, Ilhyock Shim, Hyun Song Shin, and Nikola Tarashev as well as seminar participants at the Bank for International Settlements, Bank of England, Bank of Japan, European Systemic Risk Board, South African Reserve Bank, and EEA-ESEM Congress 2021. We are grateful to Giulio Cornelli, Diego Urbina, and Alan Villegas for expert research assistance and thank staff at the Luxembourg financial supervisory authority (CSSF) and members of the Association of the Luxembourg Fund Industry's swing pricing working group for helpful discussions. A previous version of this paper was entitled "Is the Price Right? Swing Pricing and Investor Redemptions." Corresponding author (Ulf Lewrick) e-mail: ulf.lewrick@bis.org; Tel.: +41 61 280 94 58.

1. Introduction

The outbreak of the COVID-19 crisis laid bare the systemic risks that can emanate from the shadow-banking sector. In March 2020, open-end bond funds experienced rapidly accelerating outflows amid bouts of market illiquidity (e.g., Falato, Goldstein, and Hortaçsu 2021). U.S. bond funds, for example, suffered monthly outflows of more than 5 percent of total net assets (TNA), twice the amount observed during the peak of the Great Financial Crisis (GFC) of 2007/08.[1] Central banks intervened at unprecedented scale to stabilize markets and to prevent runs on funds, reminiscent of the safety net they had provided during the GFC.

Fund managers are equipped with a variety of tools to manage the risk of large-scale redemptions (International Organization of Securities Commissions (IOSCO) 2015). Yet despite the GFC experience, risk-management tools in the fund industry remain largely predicated on a microprudential approach. Relatively little is known about the effectiveness of these tools in supporting financial stability, highlighting the need for policy review (e.g., Financial Stability Board 2020).

Systemic concerns focus on the risks associated with open-end bond funds' liquidity mismatch: these funds invest in assets, such as corporate bonds, that often become illiquid under stressed market conditions, while granting fund investors the right to redeem their shares for cash daily. Concerted investor redemptions can thus force these funds to fire-sell assets. This can trigger adverse spillovers on the valuations and functioning of the underlying bond markets (e.g., Jiang, Li, and Wang 2021; Ma, Xiao, and Zeng 2020) and, given the growing role of bonds as a source of corporate funding, weigh on firms' funding conditions.

In this paper, we assess the effectiveness of swing pricing through the lens of a macroprudential perspective.[2] Swing pricing has several

[1]European bond funds, by comparison, experienced outflows equivalent to 4.9 percent of TNA in March 2020, slightly above the 4.6 percent recorded during the height of the GFC.

[2]Liquidity buffers, redemption gates, or suspensions as well as supervisory stress tests (if accompanied by corrective supervisory action) could potentially also serve as macroprudential tools to avoid fire sales and preserve investor confidence during episodes of market stress (Cominetta et al. 2018).

features that are conducive to supporting financial stability. It allows the fund manager to adjust by the “swing factor” the fund’s net asset value (NAV) per share to reflect the estimated costs associated with investor redemptions and subscriptions, such as selling or buying assets in order to meet redemption requests or invest cash inflows. Importantly, these costs can suddenly spike during episodes of market stress (e.g., Friewald, Jankowitsch, and Subrahmanyam 2012; Kargar et al. 2021; O’Hara and Zhou 2021). It is often during these times that fund investors redeem their shares to meet their liquidity needs. Absent swing pricing or other redemption charges that are credited to the fund, the cost of liquidating assets to satisfy redemptions will only be borne by those investors that stay with the fund. Anticipating this dilution of the NAV, investors have a first-mover advantage in withdrawing from the fund—creating the breeding ground for a run on the fund and destabilizing fire sales.

We present a global game that formalizes how swing pricing affects investor behavior. We illustrate how the fund can mitigate the first-mover advantage by passing the dilution costs on to withdrawing investors. However, we also show that swing factors which reflect the liquidity costs that prevail during normal market conditions fail to offset the first-mover advantage during periods of market stress. This is despite the fact that swing pricing raises measured fund returns by reducing fund dilution.

We empirically test the predictions from the global game based on a comparative analysis of nearly 2,000 open-end bond funds. We exploit the fact that swing pricing was available to funds domiciled in Luxembourg during our period of observation (2012–17), whereas U.S. funds were not yet permitted to apply swing pricing.[3]

Consistent with the predictions derived from the global game, but in contrast to previous research, we show that swing pricing did not curb outflows during the 2013 U.S. “taper tantrum.” This episode represents an ideal test case since it was characterized by a sharp but short-lived decline in fund returns that was largely contained to funds investing in fixed-income instruments. It thus did

[3]U.S. open-end funds are allowed to apply swing pricing since November 2018. However, the institutional structure of the U.S. market and operational challenges have prevented the adoption of swing pricing by U.S. funds to date (Kashyap, Kohn, and Wessel 2021).

not prompt market intervention by the public sector at any scale comparable to the one observed during the GFC or COVID-19 crisis that would blur the assessment.

We argue that swing pricing rules, which tend to be based on applying a constant swing factor once outflows exceed a certain threshold, fail to offset investor first-mover advantages in stressed markets. Because the cost that funds charge investors for their liquidity provision rises only modestly, the funds remain prone to runs in the advent of aggregate liquidity shocks. Even so, swing pricing funds exhibited higher returns during the taper tantrum. Their market-adjusted returns exceeded those of their peers by about 17 to 36 basis points (annualized) on average, at a time when the average fund return fell to −19 basis points below the funds' benchmark returns. This tallies with our model's prediction that swing pricing contains the dilution of the fund value.

Swing pricing does benefit funds when shocks are idiosyncratic. Flows of swing pricing funds are less sensitive to negative returns during normal times—as predicted by the impact of swing pricing on investor incentives. Specifically, outflows are reduced by about 0.07 percent of TNA for every percentage-point decline in returns. A swing pricing fund that exhibits a one-standard-deviation decline in (negative) returns thus benefits from a reduction in outflows of roughly 1 percent of TNA, equivalent to about half the cash holdings of the median U.S. fund in our sample. This difference arises only if returns are negative. By contrast, positive returns do not stimulate meaningful inflows into either swing pricing funds or their peer funds. The difference is most pronounced for funds investing in relatively illiquid assets, whereas it dissipates when comparing funds that invest in highly liquid securities.

Our results prove robust to testing a variety of different return measures, accounting for differences in investors' units of account and benchmark returns, as well as to considering changes in funds' market shares as an alternative approximation of fund flows.

Overall, swing pricing could provide a useful macroprudential tool to bolster the resilience of funds. However, adjustments to swing pricing rules appear necessary to ensure that investor first-mover advantages are mitigated during episodes of stress. In addition, macroprudential authorities may wish to consider topping up swing factors of large funds or those with common exposures to make these funds internalize the adverse price impact of asset liquidations in response to large redemptions.

The rest of the paper is organized in four sections. In Section 2, we discuss the related literature. Section 3 presents the rationale for swing pricing and develops a global game allowing us to derive a number of testable predictions. We apply these predictions to the data in Section 4, where we assess the impact of swing pricing on several fund performance measures, highlighting the differences between the impact under normal market conditions and during the 2013 taper tantrum. Section 5 concludes.

2. Related Literature

Our paper is related to a growing strand of the literature that studies financial stability risks arising from mutual funds and the tools to manage such risks. Jin et al. (2022) study the role of alternative pricing schemes, such as dual pricing and swing pricing, in dampening fund outflows. Based on data for about 230 U.K.-oriented funds, they find that alternative pricing schemes can dampen investor outflows including during periods of stress. This stands in contrast to our finding for a much larger sample of Luxembourg funds during the taper tantrum. One possible reason for this discrepancy is the steep downward adjustment of the NAV observed for U.K.-oriented funds during the GFC in the analysis of Jin et al. (2022), which has likely contributed to dampening fund outflows during this episode. For Luxembourg funds—similar to the approach taken in the United States—regulation limits funds' discretion to lower the NAV. This implies that first-mover advantages can surface during episodes of market stress, as also suggested by Malik and Lindner (2017), who analyze samples of up to six individual funds.[4]

Capponi, Glasserman, and Weber (2020) propose a model in which informed fund investors can anticipate redemptions and the resulting dilution of the fund value, creating a first-mover advantage. Swing pricing can offset this advantage in their model by accounting

[4] A notable feature of the data set in Jin et al. (2022) is the availability of confidential supervisory information on the funds' swing pricing activity. Our analysis, by comparison, relies on assessing investor behavior based on information that is available to investors at the point of deciding whether to redeem their shares.

for the dilution in the settlement price. This requires the swing factor to increase for larger redemptions. Modest swing factors may thus fail to contain outflows during stress episodes, consistent with our empirical findings.

Lewrick and Schanz (2017) derive welfare-optimal swing pricing policies in a general equilibrium framework. They show that trading frictions and investors' liquidity needs determine the fund manager's ability to swing the settlement price in the presence of no-arbitrage conditions. Less liquid markets would thus allow for a more active use of swing pricing. This resonates with our finding that funds investing in illiquid bonds benefit relatively more from swing pricing than those investing in more liquid bonds.

Our work also builds on that of Chen, Goldstein, and Jiang (2010), who develop a model showing how costly redemptions dilute a fund's NAV per share, creating an incentive for investors to run on the fund. They provide evidence that equity funds investing in less liquid assets experience greater outflows in response to poor performance. Goldstein, Jiang, and Ng (2017) show that this effect is even more pronounced for corporate bond funds, given the higher cost of liquidating the underlying assets.[5] Furthermore, their results imply a concave shape of the flow-to-performance relation for corporate bond funds: bond fund outflows are more sensitive to bad performance than inflows are to good performance—in contrast to the case of equity funds. This relation points to the risk of self-reinforcing redemptions during periods of weak fund performance. Our findings confirm the concave flow-to-performance relation for U.S. bond funds, while suggesting that flows for swing pricing funds are less susceptible to bad performance.

Aramonte, Scotti, and Zer (2020) study the liquidity profile of funds based on the sensitivity of fund returns to aggregate liquidity shocks. They show that less liquid funds are more exposed to redemptions in response to adverse macroeconomic news. This result tallies with our finding that funds with less liquid portfolios benefit

[5] A few empirical studies consider the dilution of fund value that arises from fund flows. Edelen (1999) finds that liquidity-motivated trading has a significant adverse effect on open-end U.S. equity fund performance. Greene and Hodges (2002) confirm this result for international funds but find no significant effect for other fund categories.

more from the dampening effect of swing pricing on fund outflows at least during normal market conditions.

3. How Does Swing Pricing Affect Investors' Behavior?

3.1 Swing Pricing in Practice

Flows in and out of an investment fund dilute the fund's value whenever they prompt the fund manager to trade securities. This is due to liquidation costs, which comprise direct transaction costs (e.g., commissions, fees) and, if the transactions are large enough, the cost due to the adverse impact on market prices.

The dilution of the fund's value creates an externality that makes funds vulnerable to the risk of runs. When investors redeem their shares, they receive the value of the fund's NAV per share. The NAV is fixed on the day the order is placed and does not incorporate the cost of fulfilling the investors' orders. Since this cost is exclusively borne by the investors that remain with the fund, there is an advantage in being among the first investors to redeem.

Swing pricing aims to reduce this externality. In principle, the externality could be eliminated by allocating to the orders the costs of their fulfillment. However, in practice, fund managers do not know these costs when fixing the NAV. Instead, fund managers adjust the NAV per share (p) by the *swing factor* (s), which approximates these costs. The "swung" NAV per share, $\tilde{p}$, at which all orders are subsequently settled, is then equal to $\tilde{p} = (1 - s)\, p$.

Various implementations of swing pricing trade off the desire to reduce the dilution of the fund's value with the need for operationally efficient and transparent rules of application. Funds typically set the swing factor equal to the approximate cost of selling securities under normal market conditions. While these factors are periodically reviewed and adjusted, there are limits to how quickly and by how much they can be raised in crisis times.[6] In addition, most

[6]According to the Association of the Luxembourg Fund Industry (ALFI) (2015), around half of the fixed-income funds cap the swing factor at a maximum of 2 percent and around half of all funds review their swing factor only at a quarterly frequency. This tallies with the description of the swing pricing policy in the prospectus of several major fund-management companies studied in our analysis. U.S. regulation, effective November 2018, applies a maximum swing factor of 2 percent of a fund's NAV per share.

funds apply a *partial* swing pricing policy, in which the swing factor is positive (negative) only if total net outflows (inflows) exceed a specified threshold (ALFI 2015).

3.2 A Model of the Effect of Swing Pricing on Investors

We present a global game to develop three testable hypotheses on how swing pricing affects investor behavior. Our model builds on the one developed in Chen, Goldstein, and Jiang (2010). We assume there is one fund and a continuum of risk-neutral investors, $[0, 1]$. Each investor initially holds one share of the fund. We normalize the total amount invested to 1. There are two periods: 1 and 2. In period 1, a fraction $\bar{X}$ of the investors decide whether to redeem their shares or whether to stay invested in the fund until period 2, when the fund is closed. All other investors remain with the fund and there are no inflows. To service redemptions, the fund needs to sell $(1+\lambda)$ worth of securities to raise one unit of cash, with $\lambda > 0$ representing the liquidation costs. Following Chen, Goldstein, and Jiang (2010), we assume the fraction $\bar{X}$ is sufficiently small to rule out that the fund has to pay out more than the amount of available funds.

Investors' actions give rise to strategic complementarities, which can bring about multiple equilibria. Each investor compares the return from withdrawing in period 1 with that from remaining invested until period 2. If she withdraws, she receives the NAV per share R_1 but will be charged the swing factor $s > 0$.[7] We assume without loss of generality that $R_1 = 1$ such that her payoff is equal to $(1 - s)$. If she stays with the fund, her payoff depends on the return of the portfolio in period 2, R_2, and the share of redeeming investors, x. The larger the redemptions, the higher the dilution and hence the lower the return from remaining with the fund, and the more optimistic about R_2 the investor has to be to stay.

To ensure a unique equilibrium, we follow the global games literature and make additional assumptions about the fund's return and investors' information.[8] We assume that R_2 increases in the random fund fundamental θ, drawn from a uniform distribution on the real

[7]This is in line with the fact that the vast majority of funds in Luxembourg only use a simple swing pricing rule, according to which a single swing factor is applied once net flows exceed the threshold.

[8]See Morris and Shin (2003) and the literature reviewed therein.

line. Each investor i receives a private signal $\theta_i = \theta + \sigma\,\varepsilon_i$ in period 1 about the unobserved fund fundamental. ε_i is an idiosyncratic noise term drawn from the distribution $g(\cdot)$, with cumulative distribution function $G(\cdot)$. The parameter $\sigma > 0$, in turn, indicates the noisiness of the signal.

In the resulting equilibrium, all investors apply a threshold strategy. If the private signal is below the threshold θ^*, the investor withdraws from the fund, whereas she remains invested if her signal is above θ^*. Morris and Shin (2003) present the proof for the general case of a symmetric binary action global game like the one presented here. For the sake of brevity, we do not restate their proof but apply their result to our case.[9]

The threshold signal θ^* is given by

$$R_2(\theta^*) = (1-s)\left[\int_{x=0}^{\bar{X}} \frac{1-(1+\lambda)\,(1-s)\,x}{1-x}dx\right]^{-1}. \tag{1}$$

To see this, notice that the investor is indifferent between remaining invested in the fund and withdrawing in period 1 if the expected return on her fund share is equal to the return she yields from redeeming her share:

$$ER_2\frac{1-(1+\lambda)\,(1-s)\,x}{1-x} = 1-s. \tag{2}$$

Rewriting the indifference condition in (2) in terms of the investor's expected payoff gain, it must hold that for the investor receiving signal $\theta_i = \theta^*$:

$$\pi(x,\theta^*) = \int_{\theta=-\infty}^{\infty} \frac{1-(1+\lambda)(1-s)x}{1-x} R_2(\theta)\frac{1}{\sigma}g\left(\frac{\theta^*-\theta}{\sigma}\right)d\,\theta - (1-s) = 0. \tag{3}$$

[9] Ensuring uniqueness of the equilibrium requires some restrictions on the swing pricing rule, $s(\cdot)$, if we allow the rule to depend on investor outflows. In particular, the rule needs to preserve action monotonicity so that the investor's expected payoff gain from remaining invested as opposed to withdrawing, $\pi(x,\theta_i)$, is non-increasing in x. The incentive to withdraw from the fund thus (weakly) increases with the share of investors that also decide to withdraw. This ensures that the rule does not reverse the first-mover advantage, which is implied by the effect of dilution, by making it more profitable for an investor to remain with the fund when an increasing share of investors decides to withdraw.

Here, $(1/\sigma)g\left((\theta^* - \theta)/\sigma\right)$ is the posterior distribution of θ conditional on having received the signal θ^*, that is, the signal of the investor who is exactly indifferent between withdrawing and remaining with the fund. Since $x = G((\theta^* - \theta)/\sigma)\bar{X}$ is the proportion of investors who withdraw, the above indifference constraint can be written as

$$\pi(x, \theta^*) = \int_{x=0}^{\bar{X}} \frac{1-(1+\lambda)(1-s)x}{1-x} R_2\left(\theta^* - \sigma\, G^{-1}\left(\frac{x}{\bar{X}}\right)\right) dx - (1-s) = 0, \tag{4}$$

which implicitly characterizes the threshold signal θ^*. In the limiting case when the investor signal becomes increasingly precise, (4) converges to (1). Equation (1) motivates three hypotheses that we test in our empirical analysis.

First, we consider how swing pricing affects the sensitivity of investor redemptions to weak fund performance. In equilibrium, all investors with signals below θ^* redeem their shares. A positive swing factor (i.e., a downward adjustment of the NAV) reduces the cutoff signal and, as a result, reduces the fraction of investors who withdraw from the fund (i.e., lowers x). This is due to the decline in the payout to redeeming investors and the increase in the expected payoff for those who remain invested.

Past returns provide a useful proxy of investor signals. Individual investor signals (θ_i) are generally not observable in practice. However, both theoretical (e.g., Berk and Green 2004; Franzoni and Schmalz 2017) and empirical research (e.g., Ben-David et al. 2022; Ben-Rephael 2017) have underscored the role of past returns in predicting investor behavior. Our analysis thus builds on using past fund returns as a gauge of investors' return expectations, in line with the literature studying the relation of fund flows and past returns.[10] Applied to our setup, we conjecture the following:

HYPOTHESIS 1. *Swing pricing reduces fund outflows in response to weak fund performance.*

[10] See Chevalier and Ellison (1997) or Sirri and Tufano (1998) for early contributions to this literature.

For our empirical analysis, this implies that swing pricing funds exhibit lower outflows than comparable U.S. funds if the investors receive a moderately weak signal. For sufficiently strong signals, by contrast, the flow-to-performance relation would be expected to be comparable across swing pricing funds and their peers, given that the expected returns of remaining invested would exceed those from withdrawing for most investors regardless of whether the fund applies swing pricing.

Our second hypothesis is concerned with the effectiveness of swing pricing as a financial stability tool given its current application. Since liquidation costs (λ) rise under stressed market conditions, funds would need to raise their swing factors in response to rising liquidation costs and increasing outflows to offset the first-mover advantage.[11]

In practice, however, swing pricing rules tend to be simple, with swing factors confined to relatively low values, subject to an upper bound and reviewed only periodically. This has important consequences for the usefulness of swing pricing as a financial stability tool. We thus conjecture the following:

HYPOTHESIS 2. *Swing factors that are calibrated to normal market conditions fail to offset investor first-mover advantages during periods of market stress.*

In this context, a key consideration relates to the link between asset sales and liquidation costs. For large funds or concerted sales by funds with common exposures, the liquidation of assets risks raising endogenously the associated liquidation costs (i.e., turning λ into an increasing function of x). This would rationalize a macroprudential top-up of the swing factor in order to account for the negative externality that the funds' sales impose on market conditions.

Our third prediction relates to the effect of swing pricing on the fund's returns:

HYPOTHESIS 3. *Swing pricing raises measured fund returns by reducing fund dilution, particularly during periods of large fund outflows.*

[11]Specifically, the fund would need to set the swing factor to $s = \lambda x/(1 + \lambda x)$ such that $\theta^* = R_2^{-1}(1)$ to ensure that investors remain invested unless they expect R_2 to drop below R_1.

This is due to two effects. First, for a given amount of outflows, the fund reduces the payout to redeeming investors by adjusting the NAV downwards (preserving sx in the fund). In addition, the fund incurs lower liquidation costs since it needs to sell fewer securities to service redemptions (preserving $sx\lambda$).[12] While we expect swing pricing to support fund returns whenever the fund swings the NAV, the effect should be strongest when liquidation costs are high such as during stressed periods.

4. Empirical Analysis

4.1 Data

Our analysis builds on 1,878 mutual bond funds for which we gather monthly data from Refinitiv Lipper and daily data on funds' NAV from Bloomberg. To construct our sample, we first select all actively managed open-end mutual bond funds available from Refinitiv Lipper that are registered in the United States or Luxembourg. These countries host the two largest mutual fund industries worldwide. Since swing pricing is applied at the level of the fund, we perform our analysis based on fund-level data, rather than using data at the level of individual fund share classes. Data coverage for Luxembourg funds improves significantly as of 2012, which is why we base our analysis on the period from January 2012 to April 2017. U.S. funds were allowed to use swing pricing only as of November 2018. Their fund-flow relationship is thus not affected by swing pricing during the period of observation.

The funds in the sample need to meet three criteria. First, based on the Refinitiv Lipper fund classification (henceforth referred to as the fund *style*), we keep only funds that can be allocated to a style for which we observe both U.S. and Luxembourg funds. This is to ensure that we are comparing funds with similar investment focus. Second, for each Luxembourg fund, we manually screen the management company's prospectus to include only those funds that have the

[12]The results in BlackRock (2016) indicate that such effects can be sizable. For the year 2015, the study finds an increase in annual emerging market fund returns by up to 77 basis points with funds swinging on up to 46 days.

ability to apply swing pricing.[13] Finally, we exclude all funds that invest mainly in advanced-economy sovereign debt. For these funds liquidation costs are low and, as a result, the first-mover advantage is small, as we will show in our analysis. Overall, our survivorship bias-free sample consists of 1,233 U.S. funds and 645 Luxembourg funds, split across 10 different styles such as *Bond USD Corporates* or *Bond USD High Yield*.

Throughout our empirical analysis, we control for time-varying characteristics of U.S. and Luxembourg funds as reported in Table 1. We consider several different measures of fund returns. This includes the funds' nominal returns and the funds' *alpha*, which for comparability we estimate based on a two-factor model as in Goldstein, Jiang, and Ng (2017).[14] About 60 percent of the funds report a benchmark index to evaluate their performance. These funds tightly manage their performance against the benchmark, as is evident from the low standard deviation and small range of the reported *market-adjusted returns*, which are given by the difference between the funds' returns and those of the benchmark. We will exploit this fact when comparing the performance of funds during the taper tantrum (see Section 4.5 below).

We measure funds' liquidity by benchmarking their cash ratio against the corresponding value of comparable funds. Specifically, we calculate an indicator variable, which is equal to one (zero otherwise) for funds with a cash ratio below the median of all funds allocated to the same style. This takes into account that funds' use of derivatives (Vivar, Wedow, and Weistroffer 2020) and holdings

[13]While the majority of Luxembourg bond funds can apply swing pricing, some funds do not apply this tool. This could reflect a preference to cater to investors that trade more frequently or diminishing marginal returns to switching to swing pricing when a large number of funds has already introduced it (e.g., Capponi, Glasserman, and Weber 2020; Jin et al. 2022). We note that the exclusion of these funds does not materially affect our main results.

[14]To estimate the fund alphas, we regress daily fund excess returns on excess aggregate bond market and aggregate stock market returns, using the Vanguard Total Bond Market Index Fund return and the Center for Research in Security Prices (CRSP) value-weighted market return as proxies, respectively. Fund alphas are then calculated as the average of the intercepts of rolling-window regressions for each fund over the past year. Our results prove robust to selecting alternative estimates of alpha, such as those resulting from a standard Fama-French three-factor model.

Table 1. Summary Statistics

	N	P10	P25	P50	P75	P90	St. Dev.	Mean
U.S. Funds								
Flow (%)	62,459	−3.48	−1.33	−0.04	1.42	4.32	6.27	0.41
Total Net Assets (Log US$ Million)	62,459	3.45	4.65	6.04	7.32	8.47	1.94	5.99
Age (Log Years)	62,459	1.10	1.79	2.64	3.09	3.33	0.86	2.41
Expense Ratio (%)	62,459	0.30	0.50	0.72	0.94	1.15	0.35	0.74
Three-Month Return (%)	62,459	−7.33	−0.58	3.22	8.58	14.59	10.04	3.82
Two-Factor Alpha (%)	48,277	−5.14	−2.31	−0.94	0.32	3.43	4.16	−0.84
Market-Adjusted Return (%)	51,691	−0.47	−0.15	0.02	0.21	0.56	0.66	0.03
Cash Relative to Peers (Binary)	47,656	0	0	1	1	1	0.50	0.52
Liquidity Beta	61,466	−0.20	−0.08	−0.01	0.04	0.15	0.20	−0.02
Redemption Charges (Binary)	62,459	0	0	0	0	1	0.31	0.11
Retail (Binary)	56,862	0	0	1	1	1	0.49	0.60
Luxembourg Funds								
Flow (%)	26,040	−4.96	−1.94	−0.02	2.08	7.21	8.44	0.87
Total Net Assets (Log US$ Million)	26,040	2.88	3.98	5.13	6.28	7.51	1.96	5.10
Age (Log Years)	26,040	0.69	1.10	1.79	2.48	3.04	0.87	1.83
Expense Ratio (%)	26,040	0.58	0.95	1.25	1.53	1.83	0.55	1.25
Three-Month Return (%)	26,040	−16.52	−6.24	2.19	9.98	19.73	14.89	2.12
Two-Factor Alpha (%)	17,109	−8.35	−3.72	−0.52	2.24	6.72	6.32	−0.68
Market-Adjusted Return (%)	12,878	−0.78	−0.30	−0.03	0.25	0.73	0.92	−0.04
Cash Relative to Peers (Binary)	16,079	0	0	0	1	1	0.50	0.47
Liquidity Beta	24,492	−0.34	−0.13	0.00	0.13	0.32	0.32	−0.01
Redemption Charges (Binary)	26,040	0	0	0	0	0	0.21	0.04
Retail (Binary)	25,238	0	1	1	1	1	0.35	0.86

Note: Unbalanced sample comprising 1,233 U.S. and 645 Luxembourg mutual open-end bond funds for the period from January 2012 to April 2017 with up to 88,499 month-fund observations per variable. The table reports the number of observations (N), the 10th (P10), 25th (P25), 50th (P50), 75th (P75), and 90th (P90) percentile as well as the standard deviation (St. Dev.) and the mean (Mean) by fund domicile. Flow: monthly net fund inflows as a percentage of total net assets (TNA); Total Net Assets: natural log of TNA in millions of U.S. dollars (USD) at the fund level; Age: natural log of years since launch of the fund; Expense Ratio: monthly expense ratio as a percentage of TNA (annualized); Three-Month Return: compound return over past three months in USD terms, in percent (annualized); Two-Factor Alpha: fund alpha, in percent, estimated over the past 12 months using a two-factor model which adds excess aggregate bond market returns to the standard Fama-French one-factor model; Market-Adjusted Returns: difference between monthly returns and the corresponding return of the fund's benchmark, in percent (annualized); Cash Relative to Peers: indicator variable that is equal to one (zero otherwise) if the fund reports cash holdings below the median of funds in the same fund style category; Liquidity Beta: Monthly average of the estimated sensitivity of the funds' net asset value per share to changes in aggregate liquidity conditions. A higher beta indicates a more sensitive liquidity profile. See footnote 15 for details; Redemption Charges: indicator variable equal to one (zero otherwise) if redemption charges apply to the fund's main share class; Retail: indicator variable that equals one (zero otherwise) if the fund's main share class has a minimum subscription amount of less than or equal to 50,000 USD.

of cash-like substitutes (Chernenko and Sunderam 2016, 2020) can blur measured cash positions.

We also estimate the funds' sensitivity to changes in aggregate liquidity conditions ("liquidity beta"), building on the approach outlined in Aramonte, Scotti, and Zer (2020).[15] This approach provides an approximation of changes in funds' liquidity risk profile at higher frequency than what can typically be constructed from reported fund holdings and is available for a large number of funds. By relying on aggregate factors, the estimates are also more robust to noisy or stale liquidity measures of infrequently traded securities, a salient feature of corporate bonds (e.g., Goldstein and Hotchkiss 2020).

Only about 1 out of 10 U.S. funds applies redemption charges, which is consistent with the results in Chernenko and Sunderam (2016) or Goldstein, Jiang, and Ng (2017). Luxembourg funds make even less use of redemption charges. This is not surprising, given that swing pricing already provides these funds with a means of levying liquidation costs on redeeming investors.

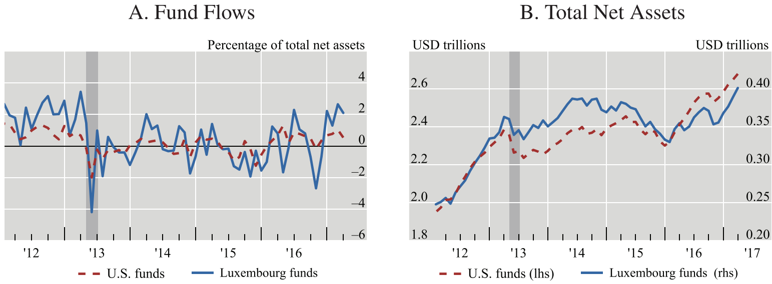
Fund flows and other fund characteristics are similar for the U.S. and Luxembourg funds, given the similarity of their investment strategies and investor basis. Figure 1A depicts the aggregate fund flows as a percentage of TNA, and Figure 1B shows TNA by country. At this level, fund flows of U.S. and Luxembourg funds as well as their TNA are highly correlated, with correlation coefficients of 0.72 and 0.85, respectively.[16]

---

[15] We estimate the liquidity profile for each fund individually based on a rolling regression over 90-day windows for funds with at least 30 observations per window: $Ret_{i,t} = \alpha_i + \beta_{Li} Liq_t + \gamma_{Zi} Z_t + \gamma_{Xi} X_{i,t} + \delta_y + \varepsilon_{i,t}$, where $Ret_{i,t}$ is the daily return of fund i, measured as the daily NAV log-changes, in excess of the return on three-month U.S. Treasury bills, our proxy of the risk-free rate. Liq_t is the negative of the noise measure proposed by Hu, Pan, and Wang (2013), such that higher values imply better aggregate liquidity conditions in bond markets. Funds with a higher liquidity beta (β_{Li}) are more sensitive to changes in aggregate liquidity risk and thus have a more risky liquidity profile. As in Aramonte, Scotti, and Zer (2020), we include in Z_t controls for changes in the level and slope of the U.S. yield curve as well as the investment grade and high-yield Markit CDX spreads, respectively. $X_{i,t}$ controls for the fund's log net asset value and log age (in years plus 1), whereas α_i and δ_y represent the constant term and year fixed effects.

[16] Flows are calculated in the standard way: $Flow_{i,t} = [TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t})]/TNA_{i,t-1}$, with $R_{i,t}$ equal to fund i's nominal return in month t. We winsorize flows at 1 percent.

Figure 1. Aggregate Fund Flows and Net Assets by Country



Note: Unbalanced sample comprising 1,233 U.S. funds and 645 Luxembourg funds. The gray-shaded region indicates the period from May 2 to July 5, 2013, when yields on longer-term U.S. Treasury securities rose sharply in reaction to policy statements by the U.S. Federal Reserve (U.S. "taper tantrum").

The period from May to early July 2013, highlighted by the gray-shaded region, marks a clear break in the flow patterns for both countries. This period was characterized by a sharp increase in bond yields, following signs of a possible tapering of the U.S. Federal Reserve's monetary policy accommodation. During this "taper tantrum," credit spreads on corporate bonds and emerging market economy debt increased significantly, resulting in sizable valuation losses for the type of funds in our sample (Bank for International Settlements 2013).

This episode coincided with a marked decline in broker-dealers' commitment of capital to support liquidity in corporate bond markets (e.g., Bessembinder et al. 2018), which amplifies illiquidity in times of stress (e.g., Bao, O'Hara, and Zhou 2018). This decline also implies an increase in search costs, particularly for larger trades, and in the implicit costs of desired trades that could not be executed. Conventional measures of liquidity, such as bid-ask spreads or the cost of executed trades, may fail to take account of this shift in broker-dealers' business models and thus overstate bond liquidity (Goldstein and Hotchkiss 2020), particularly for larger redemption-induced trades of bond funds. While, for instance, average bid-ask spreads in U.S. corporate bond markets ticked up only modestly during the taper tantrum, Dannhauser and Hoseinzade (2022) document a steep increase in the discount on the price of less liquid bond exchange traded funds, indicative of strains on dealers' intermediation capacity and tight liquidity conditions at the time.

Unlike the GFC or the COVID-19 crisis, the taper tantrum did not result in massive public sector intervention. This episode thus provides an ideal test case of the ability of swing pricing to mitigate redemption pressures absent a public sector backstop.

4.2 Methodological Approach

Our methodological approach builds on the premise that investors subscribe to funds or redeem their shares based on the information they have about a fund at any given point in time. In this sense, we presume that investors in Luxembourg funds are aware of the risk of the NAV being swung and adjust their trading behavior accordingly. This information, often supplemented by a commitment to a maximum swing factor, is available from the fund's prospectus. However, funds do not disclose the threshold nor whether they swung the NAV. For the evaluation of the impact of swing pricing on investor incentives and the corresponding investment decisions, the fact that the fund *can* swing thus appears more relevant than whether the fund *actually* swung the NAV in any given period.

Our empirical analysis is also based on the fact that, with the exception of swing pricing, U.S. and Luxembourg fund managers could resort to the same set of policy tools to address redemption pressures during the period of observation (IOSCO 2015).[17] At the same time, it appears unlikely that the ability to swing prices has any meaningful impact on the fund company's decision whether to register a fund in the United States or in Luxembourg. Other considerations, such as having established a renowned brand name in the region, are likely to dominate. Thus, we can consider the ability of the fund to swing as largely exogenous. Controlling for other fund characteristics, we can therefore gauge whether differences in the fund performance of U.S. and Luxembourg funds are consistent with the predictions from the model.

[17]We note that Luxembourg funds are, in principle, also allowed to charge anti-dilution levies on an individual transaction basis. These levies are applied to large orders of individual clients. They are thus less relevant for the investor coordination problem studied in our paper.

4.3 Differences in the Flow-to-Performance Relation of Swing Pricing Funds and Their Peers

We start with testing Hypothesis 1 by analyzing whether the flows of funds that can apply swing pricing are less sensitive to weak performance than those of their peers. We run the following regression to estimate the flow-to-performance relation:

$$\begin{aligned} Flow_{i,t} = \alpha_i &+ \beta_1\, R_{i,t-1} + \beta_2\, NegR_{i,t-1} + \beta_3\, (SP_i \times R_{i,t-1}) \\ &+ \beta_4\, (SP_i \times NegR_{i,t-1}) + \gamma_1\, X_{i,t-1} \\ &+ \gamma_2\, (SP_i \times X_{i,t-1}) + \delta_{sct} + \varepsilon_{i,t}\,, \end{aligned} \tag{5}$$

where $Flow_{i,t}$ represents the value of fund i's net inflows as a percentage of TNA in month t. α_i controls for time-invariant individual fund effects.

$R_{i,t-1}$ is the fund's lagged performance. $NegR_{i,t-1}$, in turn, is equal to R_{it-1} times an indicator variable, which is equal to one (zero otherwise) if the fund's return is negative. This accounts for potential non-linearity in the flow-to-performance relation, i.e., that investor flows respond differently to weak returns than to strong ones as motivated by our model.[18]

SP_i is a binary variable with value one if the fund applies swing pricing (i.e., is domiciled in Luxembourg) and is otherwise equal to zero. Importantly, we allow for the coefficients on all observable fund characteristics to differ between U.S. and Luxembourg funds in order to account for any underlying differences in the two groups.

We consider lagged fund controls, $X_{i,t-1}$, comprising the first lag of fund flows, log TNA, log age, and the expense ratio. We also include the indicator variable that identifies funds for which cash holdings in the previous month were below the median of those reported by all funds allocated to the same style.

We saturate the regression with fund style-country-month fixed effects, captured by δ_{sct}. This is to account for, e.g., differences in investor clienteles across fund styles or tax-loss selling before year-end, which is more prevalent for U.S. funds than for those

[18]Similar parametric regressions have been considered in the literature by, for example, Goldstein, Jiang, and Ng (2017) or Vivar, Wedow, and Weistroffer (2020).

domiciled in Luxembourg. While this conservative choice of fixed effects captures a significant share of variation across funds, it provides additional confidence in the robustness of the estimated flow-to-performance relation. $\varepsilon_{i,t}$, finally, is the error term.

Our main results are based on the first lag of compound three-month returns. This performance metric is readily available to investors when taking their decision whether to sell or buy shares. It also addresses potential endogeneity concerns that would be associated with using current returns. Furthermore, relying on nominal returns is consistent with the findings in, e.g., Ben-David et al. (2022) and Fulkerson, Jordan, and Riley (2013), who make the case that investors focus on simple return measures or composite ratings, largely neglecting any risk adjustment. All that said, we consider a variety of alternative return measures to confirm the robustness of our results.

As a reference point, we run a fund fixed-effect regression based on using all funds in the sample.[19] We report the slope coefficients on returns and negative returns in column 1 of Table 2, while also depicting the results for less saturated versions of the regression in columns 2 and 3.[20] The estimates accord with previous findings in the literature of a concave flow-to-performance relation (e.g., Goldstein, Jiang, and Ng 2017), as indicated by the much larger coefficient on negative returns than the one on (all) returns for U.S. funds.

Based on column 1, an increase in a representative U.S. fund's annualized returns by 10 percentage points in the preceding three-month period (roughly equivalent to one standard deviation of the U.S. funds' returns in our sample) would lead to additional inflows of only 0.2 percent of TNA in the current month. Had the fund,

[19] Alternative estimation procedures, such as ordinary least squares (OLS) or the widely applied GMM estimator proposed by Arellano and Bond (1991), do not appear preferable to the fixed-effect regression for our purposes. The OLS estimation is expected to be biased in a dynamic panel setup, as is confirmed by a much higher coefficient estimate for lagged flows (not reported) compared with the estimate of the fixed-effect regressions. The GMM estimator, in turn, may overidentify the model given the large number of instruments that result from using a sample with up to 54 monthly observations per fund. Indeed, the length of the sample argues in favor of using the fixed-effect regression in our case.

[20] For brevity, we only report the slope coefficients on returns. All other estimates are available upon request from the authors.

Table 2. Flow-to-Performance Relation: Funds With vs. Funds Without Swing Pricing

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
R_{t-1}	0.023*	0.020*	0.022*	0.014	0.062**	0.140**	0.042
	(0.012)	(0.009)	(0.010)	(0.015)	(0.027)	(0.061)	(0.046)
$NegR_{t-1}$	0.092***	0.062***	0.041**	0.128***	0.109***	0.166	0.142
	(0.020)	(0.009)	(0.014)	(0.020)	(0.033)	(0.104)	(0.094)
$R_{t-1} \times SP$	0.012	0.016*	0.016	0.019	0.051	0.085	0.084
	(0.012)	(0.008)	(0.010)	(0.014)	(0.035)	(0.076)	(0.059)
$NegR_{t-1} \times SP$	–0.075***	–0.065***	–0.053***	–0.091***	–0.159**	–0.371**	–0.144
	(0.010)	(0.009)	(0.009)	(0.013)	(0.064)	(0.118)	(0.101)
R-squared	0.108	0.075	0.060	0.125	0.127	0.127	0.131
Number of Funds	1,878	1,878	1,878	1,479	1,467	1,442	1,128
Observations	88,499	88,499	88,499	54,965	54,785	53,886	42,735
Return Measure	3m	3m	3m	3m	6m	12m	Alpha
Cash Holdings	No	No	No	Yes	Yes	Yes	Yes
Fund Fixed Effects (FE)	Yes	No	Yes	Yes	Yes	Yes	Yes
Style × Country × Month FE	Yes	No	No	Yes	Yes	Yes	Yes
Excluding Funds with Redemption Charges	No	No	No	Yes	Yes	Yes	Yes

Note: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$. Estimates based on Equation (5) with robust standard errors, clustered by fund style, in parentheses. Dependent variable: net inflows as a percentage of total net assets (TNA). All regressions control for lagged returns, lagged negative returns, lagged log TNA, log age, and the lagged expense ratio; regressions (4) to (7) also include an indicator of cash holdings, which is equal to one if the fund's cash holdings in the previous month were above the median of corresponding cash holdings of funds with the same style. For each control variable, we include the base effect and an interaction term for swing pricing funds (SP) to allow for a different elasticity for these funds and their peer funds. Return measure (R): annualized returns (in percent) over the past 3 months (regressions (1) to (4)), 6 months (5), and 12 months (6) as well as fund alpha (7), which is based on a two-factor model, estimated as the average of the intercepts of rolling-window regressions for each fund over the past year. Negative returns (NegR): returns interacted with an indicator variable, which is equal to one (zero otherwise) if returns are negative.

by comparison, experienced a decline in returns from 0 percent to -10 percent, it would have been subject to additional net outflows of 1.2 percent of TNA ($10 \times (0.023 + 0.092)$). To put these numbers into perspective, we note that median cash holdings of U.S. funds in our sample were 2 percent of TNA.

The relation of fund flows and past returns differs for swing pricing funds and their peers, supporting Hypothesis 1. To see this, we move to the slope coefficients that result from interacting returns and the indicator for swing pricing funds (SP, column 1 of Table 2). What stands out is these funds' lower sensitivity to negative returns. The reduction in outflows that can be linked to swing pricing follows from the sum of the slope coefficients on returns and negative returns of swing pricing funds ($0.012 - 0.075$). Conditional on returns being negative, outflows are reduced by roughly 0.06 percent of TNA for every percentage-point decline in returns. A swing pricing fund with a negative three-month return dropping by 10 percentage points would thus have witnessed additional net outflows of 0.5 percent of TNA ($-10 \times (0.023 + 0.092 + 0.012 - 0.075)$) in the next month, less than half of its U.S. counterparts. Positive returns, by comparison, do not seem to stimulate meaningful inflows into either U.S. or swing pricing funds.

Controlling for additional fund characteristics that may affect investor redemptions reinforces our findings. First, we exclude funds that apply redemption charges, since such charges, similar to swing pricing, should discourage investor redemptions. Second, we account for differences in the funds' cash holdings in order to control for anticipated dilution costs. We recall from the model that higher liquidation costs increase the dilution effect of investor redemptions, reducing investor incentives to stay with the fund. Cash holdings serve as a useful gauge of liquidation costs. As a direct effect, lower (higher) cash holdings raise (reduce) the average costs of liquidating the fund portfolio. In addition, funds with lower cash holdings may have to respond more promptly to net outflows, suggesting that they have less leeway in timing their sales and may need to make higher price concessions when liquidating securities.

The effect on swing pricing funds is reinforced by applying these additional controls (column 4 of Table 2). Specifically, the marginal effect of negative returns on outflows is lowered by roughly 0.07 percent of TNA (0.019–0.091) per percentage-point of return when comparing swing pricing funds with U.S. funds. A swing pricing fund that exhibits a one-standard-deviation decline in (negative) returns thus benefits from a reduction in outflows equivalent to about 1 percent of TNA.

As an additional robustness check, we regress fund flows on the same set of predictors as before, but vary the measure of fund returns. Specifically, we consider the impact of using cumulative returns over the preceding 6 months, 12 months, and the estimated two-factor fund alpha, respectively. This adjustment assumes that investors not only factor in the recent returns to assess future fund performance, but also consider the fund's medium-term performance in their assessment. One rationale for such an approach is the presumption that skilled fund managers perform well on average, but may nevertheless fail to generate returns in individual months (Kacperczyk, van Nieuwerbaugh, and Veldkamp 2014).

The results shown in columns 5 to 7 of Table 2 lend further support to Hypothesis 1. Negative returns induce fewer outflows from swing pricing funds than from their peers. Intuitively, the difference becomes larger as we lengthen the range of returns that we assume investors are factoring into their decisions. Comparing a U.S. and swing pricing fund with negative returns over the past six months, the latter benefits from reduced outflows of about 0.11 percent of TNA (0.051–0.159) per percentage-point change in negative returns (column 5). If the return of both funds was negative over the past 12 months, the benefit amounts to as much as 0.29 percent of TNA (0.085–0.371; column 6). The results are qualitatively similar for regressions based on fund alpha, although the coefficient estimates are statistically insignificant at the usual confidence levels (column 7).

4.4 Discussion of Alternative Drivers

We assess to what extent alternative factors could be driving cross-country differences in the flow-to-performance relation. First, we assess whether differences in the flow-to-performance relation persist if we compare funds that face no meaningful dilution risks. Absent such risks, there would be no first-mover advantage and any observed differences between swing pricing funds and their peers would need to be driven by other factors.

To test this, we estimate the flow-to-performance relation for subsamples of funds with different liquidity profiles. Following Aramonte, Scotti, and Zer (2020), we measure these profiles based on the sensitivity of the funds' daily returns to aggregate liquidity factors. We expect the swing pricing effect to be strongest for funds that are most sensitive, i.e., those with the highest liquidity betas. The price of these funds declines the most when aggregate liquidity conditions worsen, consistent with their portfolio being the least liquid. Accordingly, we group the funds based on their monthly average liquidity betas, distinguishing between the most sensitive ones with betas above the 75th percentile of the sample (column 1 of Table 3), the sensitive ones with betas above the sample median (column 2), and the least sensitive ones with betas below the 25th percentile (column 3).

Differences between swing pricing funds and their peers dissipate as the liquidity risk profile of the funds improves (Table 3). The dampening effect of swing pricing on redemptions from funds with negative returns is strongest for those funds that exhibit the highest sensitivity to changes in aggregate liquidity conditions, whereas we find a much weaker and statistically insignificant effect for funds with the least sensitive profiles.

We also consider subsample regressions based on categorizing funds by their style as suggested by Chen, Goldstein, and Jiang (2010). The advantage of this approach is that it is based on a feature that is disclosed at the inception of the fund. The style is thus known to investors and is exogenous to fund flows. We estimate the flow-to-performance relation for a subsample of funds that invest in particularly illiquid bonds, e.g., emerging market bonds, high-yield bonds (column 4 of Table 3), one composed of funds investing in relatively more liquid bonds, e.g., USD short- and medium-term bonds (column 5), and a control group comprising funds that exclusively invest in the most liquid bonds, i.e., advanced-economy sovereign bonds (column 6).[21]

[21] The sample of sovereign bond funds comprises 151 U.S. funds and 94 Luxembourg funds which are not included in the main sample presented in Table 1. To ensure that the results are representative, the regression in column 3 of Table 3 does not include funds' cash holdings, for which data are missing for many funds. For this type of fund, cash holdings are likely to be less relevant given the high liquidity of the fund portfolio. Accordingly, differences in the flow-to-performance relation remain insignificant if cash holdings are accounted for in the regression.

Table 3. Flow-to-Performance Relation: The Effect of Portfolio Liquidity

	Portfolio by Liquidity Risk Profile			Portfolio by Fund Style		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
R_{t-1}	–0.010	0.020	0.016	0.014	0.018	0.010
	(0.024)	(0.024)	(0.017)	(0.024)	(0.019)	(0.019)
$NegR_{t-1}$	0.201***	0.133***	0.111**	0.162**	0.106***	0.004
	(0.061)	(0.037)	(0.038)	(0.041)	(0.016)	(0.041)
$R_{t-1} \times SP$	0.040	0.009	0.038	0.014	0.016	–0.014
	(0.030)	(0.025)	(0.032)	(0.015)	(0.016)	(0.068)
$NegR_{t-1} \times SP$	–0.205***	–0.141***	–0.059	–0.107**	–0.082***	0.061
	(0.051)	(0.040)	(0.043)	(0.023)	(0.008)	(0.103)
R-squared	0.196	0.158	0.171	0.178	0.089	0.147
Number of Funds	1,307	1,435	1,241	436	982	245
Observations	12,895	26,939	12,821	15,587	36,849	12,626
Portfolio Liquidity	Most Sensitive	Sensitive	Least Sensitive	Illiquid	Liquid	Most Liquid
Return Measure	3m	3m	3m	3m	3m	3m
Cash Holdings	Yes	Yes	Yes	Yes	Yes	No
Fund Fixed Effects (FE)	Yes	Yes	Yes	Yes	Yes	Yes
Style × Country × Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Excluding Funds with Redemption Charges	Yes	Yes	Yes	Yes	Yes	Yes

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Estimates based on Equation (5) with robust standard errors, clustered by fund style, in parentheses. Dependent variable: net inflows as a percentage of total net assets (TNA). Funds categorized by liquidity risk profile: (1) funds for which the monthly liquidity beta (estimated sensitivity to changes in aggregated liquidity conditions) is above the 75th percentile of the sample; (2) funds with monthly liquidity beta above the sample median; (3) funds with monthly liquidity beta below the 25th percentile. Funds categorized by liquidity of fund style: (4) funds investing in relatively illiquid assets (e.g., emerging market bonds, high-yield bonds); (5) funds investing in relatively liquid assets (e.g., USD short- and medium-term bonds, global bonds), (6) control sample of 245 funds investing predominantly in the most liquid assets, i.e., advanced-economy sovereign bonds. All regressions control for lagged returns, lagged negative returns, lagged log TNA, log age, and the lagged expense ratio; regressions (1) to (5) also include an indicator of cash holdings, which is equal to one if the fund's cash holdings in the previous month are above the median of corresponding cash holdings of funds with the same style. This variable is excluded in (6) given that funds in this subsample focus on most liquid assets. For each control variable, we include the base effect and an interaction term for swing pricing funds (SP) to allow for a different elasticity for these funds and their peer funds. Return measure (R): annualized returns (in percent) over the past three months. Negative returns (NegR): returns interacted with an indicator variable, which is equal to one (zero otherwise) if returns are negative.

Sovereign bond funds exhibit no notable cross-country difference in their flow-to-performance relation. For funds investing in illiquid bonds, by contrast, the difference is significant and somewhat more pronounced than for the intermediate case of funds invested in relatively more liquid bonds. These findings lend support to our interpretation that by reducing the risk of dilution, swing pricing dampens the funds' sensitivity of flows to weak performance.

Second, we test whether potential differences in investors' currency of account and opportunity costs affect our results. One concern with the above regressions could be that investors from different regions evaluate fund performance based on their local currency. Since an appreciation (depreciation) of the USD against the investor's local currency raises (reduces) the fund's return in local-currency terms, investors using different currencies of account could respond differently to past fund returns. While information on investors' currency of account is generally not available, a rough approximation is that investors are more likely to evaluate U.S. fund performance in USD terms and Luxembourg fund performance in euros (EUR), for example, because investors are biased towards investing in funds domiciled in their home region.

Another concern relates to measuring investors' opportunity costs. Thus far, we have implicitly assumed that investors respond differently to negative returns than to positive returns because a natural alternative to investing in funds—at least in the short term—is to hold cash, which yields zero nominal return. We alter this assumption and consider how our results change if we assume that investors benchmark fund returns against the corresponding risk-free rates. Specifically, we replace our identifier of negative returns in Equation (5) with one that indicates whether the fund returns fell below the risk-free rate ($BelowRF_{i,t-1}$). For Luxembourg funds, our approximation of the risk-free rates is given by the euro-currency market interest rates with the corresponding term. For U.S. funds, we use the corresponding yields on the U.S. Treasury bills.[22]

[22]Differences in the funds' investor base could influence the flow-to-performance relation. Studying equity funds, Ferreira et al. (2012) relate cross-country variation in this relation to differences in investor sophistication and participation costs. Yet such differences are small for the United States and Luxembourg, as gauged from the proxies used in their study, such as average education levels or the quality of the judicial system.

Our results prove robust to varying investors' currency of account and opportunity costs. In columns 1 and 2 of Table 4, we show the slope coefficients of interest based on measuring U.S. fund returns in USD (as throughout the paper), but swapping Luxembourg fund returns into EUR terms based on spot exchange rates. In both cases, we control for returns falling below the risk-free rate with the corresponding term. We note that the estimated effect of swing pricing is little changed when compared with the corresponding estimates in Table 2 (columns 4 and 5).

Finally, we inspect changes in funds' market shares as an alternative measure of fund flows. Spiegel and Zhang (2013) question the validity of the standard flow-to-performance specification. They argue that measuring flows as a percentage of TNA yields biased results and propose using changes in market shares to obtain robust estimates of the flow-to-performance relation. To consider this, we repeat the above analysis based on using this alternative measure of flows. Following Spiegel and Zhang (2013), we calculate the change in fund i's market share, $\Delta m_{i,t}$ as

$$\Delta m_{i,t} = \frac{TNA_{i,t}}{\sum_{j \in \Omega_{t-1}} TNA_{j,t}} - \frac{TNA_{i,t-1}}{\sum_{j \in \Omega_{t-1}} TNA_{j,t-1}}, \tag{6}$$

where Ω_{t-1} comprises all the funds with the same style as fund i that were in existence in period $t-1$. We measure changes in market share in basis points and, as for fund flows, winsorize at 1 percent.

Our finding of swing pricing funds' lower sensitivity to poor performance is robust to using this alternative measure of fund flows. Columns 3 to 5 of Table 4 report the corresponding results for different specifications, taking into account the role of cash holdings, varying investors' currency of account, and opportunity costs. In each case, we find a statistically significant reduction in the response of market shares to negative performance for swing pricing funds.

4.5 Fund Performance During the Taper Tantrum

We now turn to assessing differences in fund performance between swing pricing funds and their peers during the taper tantrum. Overall, we detect little evidence of systematic differences in net fund

Table 4. Flow-to-Performance Relation: Accounting for Exchange Rates, Opportunity Costs, and Alternative Measures of Fund Flows

	Fund Flows		Change in Market Share		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
R_{t-1}	0.014	0.062**	–0.007	–0.007	–0.007
	(0.015)	(0.026)	(0.008)	(0.008)	(0.008)
$BelowRF_{t-1}$	0.128***	0.109***	0.018	0.026**	0.026**
	(0.020)	(0.033)	(0.011)	(0.010)	(0.010)
$R_{t-1} \times SP$	0.019	0.051	0.011	0.016	0.016
	(0.014)	(0.034)	(0.009)	(0.011)	(0.011)
$BelowRF_{t-1} \times SP$	–0.094***	–0.161**	–0.016**	–0.027***	–0.028***
	(0.013)	(0.058)	(0.006)	(0.008)	(0.008)
R-squared	0.125	0.127	0.131	0.145	0.145
Number of Funds	1,479	1,467	1,577	1,479	1,479
Observations	54,965	54,965	75,210	54,965	54,965
Return Measure	3m	6m	3m	3m	3m
Risk-Free Rate	3m	6m	0	0	3m
Currency	Local	Local	USD	USD	Local
Cash Holdings	Yes	Yes	No	Yes	Yes
Fund Fixed Effects (FE)	Yes	Yes	Yes	Yes	Yes
Style × Country × Month FE	Yes	Yes	Yes	Yes	Yes
Excluding Funds with Redemption Charges	Yes	Yes	Yes	Yes	Yes

Note: $*p < 0.1$; $**p < 0.05$; $***p < 0.01$. Estimates based on Equation (5) with robust standard errors, clustered by fund style, in parentheses. Dependent variable in (1) and (2): net inflows as a percentage of total net assets (TNA); in (3) to (5): monthly change in the fund's market share in basis points as defined in Equation (6). All regressions control for lagged returns, lagged negative returns (or returns below the risk-free rate with the corresponding tenor), lagged log TNA, log age, and the lagged expense ratio; regressions (1), (2), (4), and (5) also include an indicator of cash holdings, which is equal to one if the fund's cash holdings in the previous month are above the median of corresponding cash holdings of funds with the same style. For each control variable, we include the base effect and an interaction term for swing pricing funds (SP) to allow for a different elasticity for these funds and their peer funds. Return measure (R): returns are measured as the three-month (3m) or six-month (6m) return, annualized in percent. BelowRF: returns interacted with an indicator variable, which is equal to one (zero otherwise) if returns are below the risk-free rate indicated in the respective column. For U.S. funds, risk-free rates are approximated by the yield on U.S. Treasury bills. For Luxembourg funds, risk-free rates are approximated by euro-currency market interest rates. Currency: In regressions (1), (2), and (5), Luxembourg fund returns are converted from USD into local currency (EUR) based on monthly USD/EUR spot exchange rates.

flows of these two groups of funds during this episode. This finding tallies with Hypothesis 2, which suggests that swing pricing, given its current design, is unlikely to offset investor first-mover advantages if markets are under stress. That said, our results suggest that swing pricing funds which were subject to outflows benefitted from higher market-adjusted returns than their U.S. peers. This is consistent with the anti-dilution effect of swing pricing conjectured in Hypothesis 3.

We start with Hypothesis 2 and test for systematic differences in the flow-to-performance relation of funds during the taper tantrum. We run the following regression, which considers the triple interaction of the effect of returns, swing pricing (SP), and observations during the taper tantrum ($Stress$):

$$\begin{aligned} Flow_{i,t} = \alpha_i &+ \beta_1 R_{i,t-1} + \beta_2 BelowRF_{i,t-1} \\ &+ \beta_3 (R_{i,t-1} \times Stress_t) + \beta_4 (BelowRF_{i,t-1} \times Stress_t) \\ &+ \beta_5 (R_{i,t-1} \times SP_i) + \beta_6 (BelowRF_{i,t-1} \times SP_i) \\ &+ \beta_7 (R_{i,t-1} \times SP_i \times Stress_t) \\ &+ \beta_8 (BelowRF_{i,t-1} \times SP_i \times Stress_t) \\ &+ \gamma_1 X_{i,t-1} + \gamma_2 (X_{i,t-1} \times Stress_t) \\ &+ \gamma_3 (X_{i,t-1} \times SP_i) + \gamma_4 (X_{i,t-1} \times SP_i \times Stress_t) \\ &+ \delta_{sct} + \varepsilon_{i,t}, \end{aligned} \tag{7}$$

where we control for the possibility that the elasticity of flows with respect to all other fund characteristics ($X_{i,t-1}$) may have also varied for U.S. and Luxembourg funds during the taper tantrum. Table 5 reports each of the eight β coefficients in Equation (7) for different specifications.

In line with the above analysis, we find that swing pricing dampens the net flows of funds with returns below the risk-free rate during normal market conditions ($\beta_5 + \beta_6$) for each specification. The magnitude of the effect is also comparable to our previous results.

However, consistent with Hypothesis 2, we find no systematic difference between swing pricing funds and U.S. funds during the taper tantrum. The sum of the four coefficients (β_5 to β_8) that are interacted with the swing pricing indicator (SP) is near zero

Table 5. Flow-to-Performance Relation: Swing Pricing during Stress Episodes

	(1)	(2)	(3)	(4)	(5)
R_{t-1}	0.012	0.012	–0.007	0.020	0.068**
	(0.016)	(0.016)	(0.008)	(0.016)	(0.027)
$BelowRF_{t-1}$	0.127***	0.127***	0.025**	0.092**	0.094***
	(0.023)	(0.023)	(0.011)	(0.022)	(0.024)
R_{t-1} × Stess	0.047	0.047	0.013	–0.000	0.062
	(0.071)	(0.071)	(0.013)	(0.059)	(0.066)
$BelowRF_{t-1}$ × Stress	0.073	0.073	0.013	0.095	–0.073
	(0.221)	(0.221)	(0.045)	(0.177)	(0.113)
R_{t-1} × SP(β_5)	0.022	0.022	0.017	0.016	0.016
	(0.014)	(0.014)	(0.011)	(0.016)	(0.030)
$BelowRF_{t-1}$ × SP (β_6)	–0.099***	–0.102***	–0.029**	–0.076***	–0.112**
	(0.017)	(0.017)	(0.010)	(0.018)	(0.036)
R_{t-1}× SP × Stress (β_7)	0.085	0.079	–0.023	–0.224	–0.083
	(0.338)	(0.350)	(0.046)	(0.209)	(0.118)
$BelowRF_{t-1}$ × SP × Stress(β_8)	0.005	0.017	0.048	0.314	0.270
	(0.492)	(0.500)	(0.081)	(0.323)	(0.183)
Swing Pricing Effect:					
Normal Conditions ($\beta_5 + \beta_6$)	–0.077***	–0.079***	–0.012*	–0.060***	–0.096**
Stress ($\beta_5 + \beta_6 + \beta_7 + \beta_8$)	0.013	0.017	0.012	0.029	0.091
R-squared	0.214	0.214	0.301	0.202	0.201
Number of Funds	1,472	1,472	1,472	1,532	1,513
Observations	54,915	54,915	54,944	73,451	73,045
Dependent Variable	Flows	Flows	ΔMarket Share	Flows	Flows
Return Measure	3m	3m	3m	3m	6m
Risk-Free Rate	0	3m	3m	3m	6m
Currency	USD	Local	Local	Local	Local
Liquidity Measure	Cash	Cash	Cash	Liq. Beta	Liq. Beta
Fund Fixed Effects (FE)	Yes	Yes	Yes	Yes	Yes
Style × Country × Month FE	Yes	Yes	Yes	Yes	Yes
Excluding Funds with Redemption Charges	Yes	Yes	Yes	Yes	Yes

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Estimates based on Equation (7) with robust standard errors, clustered by fund style, in parentheses. Dependent variable in (1), (2), (4), and (5): net inflows as a percentage of total net assets (TNA); in (3): monthly change in the fund's market share in basis points as defined in Equation (6). All regressions control for lagged returns, lagged negative returns (or returns below the risk-free rate with the corresponding tenor), lagged log TNA, log age, and the lagged expense ratio. In regressions (1) to (3), the control variables also comprise an indicator of cash holdings as a measure of liquidity, which is equal to one if the fund's cash holdings in the previous month are above the median of corresponding cash holdings of funds with the same style. In regressions (4) and (5), cash holdings are replaced by the lagged liquidity beta. Stress is an indicator variable, which is equal to one (zero otherwise) for observations in May and June 2013 (taper tantrum). For each control variable, we include the base effect, an interaction term for swing pricing funds (SP), an interaction term for the taper tantrum (Stress), and the triple interaction with Stress and SP to allow for different elasticities for U.S. funds and swing pricing funds during normal times and the taper tantrum, respectively. Return measure (R): returns are measured as the three-month (3m) or six-month (6m) return, annualized in percent. BelowRF: returns interacted with an indicator variable, which is equal to one (zero otherwise) if returns are below the risk-free rate indicated in the respective column (see line "Risk-Free Rate"). For U.S. funds, risk-free rates are approximated by the yield on U.S. Treasury bills. For Luxembourg funds, risk-free rates are approximated by euro-currency market interest rates. Currency: In regressions (2) to (5), Luxembourg fund returns are converted from USD into local currency (EUR) based on monthly USD/EUR spot exchange rates.

for returns measured over a three-month horizon (columns 1 to 4). It is also statistically insignificant for funds for which returns fell below the risk-free rate if measured over a six-month horizon (column 5). The increased sensitivity of fund flows to weak performance of swing pricing funds relative to U.S. funds during the taper tantrum, as captured by the coefficient β_8 on the triple interaction of $BelowRF_{i,t-1} \times SP_i \times Stress_t$, contributes to counterbalancing the dampening effect of swing pricing on investor redemptions. This tallies with the model's prediction that it would take a large swing factor to fully offset the investor first-mover advantage during stressed market conditions.

To sharpen the previous analysis, we match individual swing pricing funds with U.S. funds (excluding those that impose redemption charges) based on a variety of fund characteristics available to investors in the run-up to the taper tantrum. We estimate the average treatment effect of swing pricing funds, the *treated* funds (ATET). We consider the ATET of several different measures. The top rows of Table 6 present results for the cumulative net fund flows from May to June 2013, winsorized at 1 percent to account for outliers. The next row reports estimates for market-adjusted returns. These are calculated as the difference between each fund's annualized returns and those of the fund's benchmark from May to June 2013. We use this measure, rather than nominal returns, to account for differences in the riskiness of fund portfolios. We recall that the funds in our sample deviate little from their benchmarks (see also Table 1), suggesting that this adjustment provides a powerful control for differences across fund returns that are not related to swing pricing. Because we expect funds to swing most frequently if they experience large outflows during this period, we report in the third row the results based on considering only those funds that experience net outflows.

We apply several alternative matching algorithms to evaluate the effect of swing pricing. Columns 1 and 2 of Table 6 report the ATET using nearest-neighbor matching. We match each swing pricing fund with four U.S. funds, using the number of neighbors recommended in Abadie and Imbens (2011). For the results in column 1, we match funds based on their log age, log TNA, nominal returns, net flows, and cash ratio relative to peers in April 2013, i.e., the month preceding the taper tantrum. We also include an indicator variable for

Table 6. Average Treatment Effect of Swing Pricing Funds during the Taper Tantrum

	(1)	(2)	(3)	(4)
Net Fund Flows	–1.656	–0.430	–0.781	0.511
	(1.324)	(1.229)	(0.845)	(0.803)
Matched Swing Pricing Funds	150	135	286	277
Market-Adjusted Returns:				
All Funds	0.113	0.166*	0.223***	0.222***
	(0.077)	(0.088)	(0.073)	(0.062)
Matched Swing Pricing Funds	147	135	154	140
Market-Adjusted Returns:				
Funds with Net Outflows	0.253***	0.268***	0.358***	0.246***
	(0.094)	(0.100)	(0.102)	(0.071)
Matched Swing Pricing Funds	100	88	97	80
Matching	Nearest Neighbor	Nearest Neighbor	ρ of Daily Returns	ρ of Daily Returns

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Robust standard errors in parentheses. The table reports the average treatment effect of swing pricing funds. Net fund flows are measured as the cumulative net flows in May and June 2013 (taper tantrum); as a percentage of total net assets (TNA). Market-adjusted returns are given by the compound nominal fund return less the return of the fund's benchmark for the period from May to June 2013, annualized in percent. Funds with net outflows comprises all funds with net outflows during the period from May to June 2013. Matching: (1) Nearest-neighbor matching with four neighbors; (2) nearest-neighbor matching with four neighbors within the same fund style category. Funds are matched on April 2013 observations (i.e., the month preceding the taper tantrum) of net fund flows, (log) TNA, (log) age, and one-month returns (all normalized by their cross-sectional standard deviation) as well as an indicator variable of whether the fund had cash holdings above or below the median of those of other funds using the same benchmark in April 2013 and an indicator variable for retail funds; (3) each swing pricing fund is matched with the U.S. fund in the same fund style category with the highest pairwise correlation based on daily returns over the three months preceding the taper tantrum; (4) each swing pricing fund is matched with the four U.S. counterparts exhibiting the highest correlation; to calculate potential outcomes for swing pricing funds, we weigh the observations for the four U.S. funds by the relative value of their correlation coefficients.

retail funds to take account of potential differences in funds' investor base. All variables are normalized by their cross-sectional standard deviation. For the results in column 2, we impose as an additional constraint that funds are only matched with funds of the same style.

Swing pricing funds do not appear to have experienced smaller outflows than their U.S. counterparts during the taper tantrum. The top row of columns 1 and 2 depicts the ATET for cumulative net fund flows during the taper tantrum. If swing pricing was effective in dampening fund outflows during the taper tantrum, we would expect to observe positive coefficient estimates. Yet, consistent with Hypothesis 2, we find no statistically significant difference between the outflows of the two groups of funds.

To gain further insights, we develop an alternative matching algorithm based on the correlation of daily fund returns. This approach builds on the presumption that funds with similar portfolios should be characterized by a high correlation of their returns.

For each swing pricing fund, we calculate the pairwise correlation coefficient of daily returns with each individual U.S. fund within the same style category over the three months preceding the taper tantrum. Since funds experienced relatively steady inflows during these months (see also Figure 1A), we do not expect their measured returns to be much affected by swing pricing activity. To further increase the precision of our comparison, we keep only correlation coefficients that are based on at least 30 non-zero observations per fund during this period. Next, we match each swing pricing fund with the U.S. fund for which we observe the highest correlation. Column 3 of Table 6 presents the corresponding ATET. For robustness, we also estimate the ATET that follows from matching each swing pricing fund with the four U.S. funds of the same style that exhibit the highest correlation with this fund in column 4.[23]

Overall, the ATET for the funds' net inflows provides no evidence of a dampening effect of swing pricing fund outflows during the taper tantrum. Seen through the lens of funds' liquidity management, this suggests that swing pricing policies were too timid to offset first-mover advantages during this episode of elevated market uncertainty.

[23]The correlation coefficients for the (fourth) closest match between swing pricing funds and U.S. funds have a median value of (0.69) 0.74.

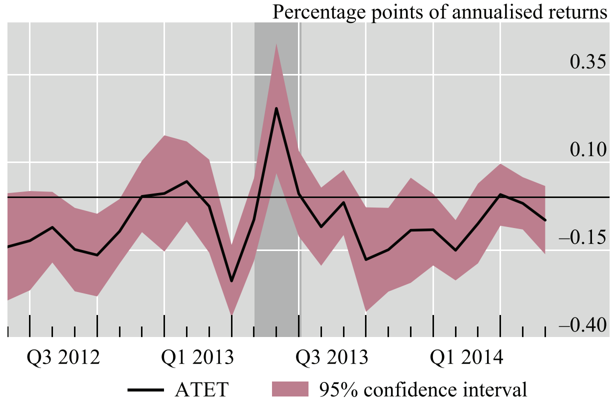
We now turn to the assessment of market-adjusted returns to test whether swing pricing funds managed to mitigate the dilution implied by fund outflows. In line with Hypothesis 3, we find evidence that swing pricing funds generated higher returns than their peers. Columns 1 and 2 report the ATET based on using nearest-neighbor matching, whereas columns 3 and 4 depict the estimates based on matching daily returns (see above). We report results based on matching all funds (middle row) and based on matching only those funds that exhibited net outflows during the taper tantrum (bottom row).

The estimates based on all funds point to additional returns in a range of about 17 to 22 basis points on an annualized basis—a sizable effect, given that the average market-adjusted return dropped to about –19 basis points during the taper tantrum (unadjusted returns averaged –21 percent) and hovered around zero when considering the entire period of observation (see Table 1). With swing pricing funds facing average net outflows of about 2.85 percent of TNA during the taper tantrum, these additional returns imply an average swing factor of about 1 percent to 1.3 percent—consistent with the values reported by the industry (ALFI 2015).

Intuitively, we find a larger effect—up to 36 basis points—if we constrain the sample to funds exhibiting outflows (Table 6, bottom row). These funds were under greater pressure to liquidate assets in order to accommodate investor redemptions. Thus, they are more likely to have incurred high liquidation costs, with only the swing pricing funds benefiting from the reduction in payouts to investors by swinging their NAV. This finding accords with Jin et al. (2022), who document that U.K.-oriented funds with alternative pricing schemes suffer less dilution due to outflows than other funds.

We test whether the difference in market-adjusted returns is not an artefact of matching generally more profitable swing pricing funds with less profitable U.S. funds. To do so, we estimate the ATET for the matched funds for each month over a two-year window centered on the taper tantrum based on comparing (annualized) market-adjusted returns for the latest two months. Figure 2 depicts the corresponding ATET and its 95 percent confidence interval using the matching approach applied for the results in column 1 of Table 6

Figure 2. Monthly ATET Estimates for Swing Pricing Funds' Market-Adjusted Return



Note: Average treatment effect of 100 matched swing pricing funds based on the regression in Table 6, column 1. Market-adjusted returns are calculated as the difference between two-month compound returns and the corresponding return on the fund's benchmark, annualized in percent. The gray-shaded region indicates the period from May 2 to July 5, 2013 (taper tantrum).

for the 100 swing pricing funds subject to net outflows during the taper tantrum. The ATET tends to hover around or slightly below zero but spikes during the taper tantrum (highlighted by the gray-shaded region), consistent with swing pricing helping these funds to contain dilution relative to their peers during this stress episode.

5. Conclusion

The COVID-19 crisis has revived concerns about systemic risks in the shadow-banking sector. The crisis thus provides a timely reminder of the need to expand the macroprudential framework to non-banks, such as open-end mutual funds. In this paper, we intend to make a first step towards developing such a framework by exploring the effects of swing pricing—a candidate tool to mitigate the risk of runs on funds and fire sales.

Based on a global game, we develop several predictions to guide our empirical analysis of how swing pricing affects investor behavior.

Our identification strategy is based on comparing Luxembourg funds that were allowed to apply swing pricing with similar funds from the United States, where swing pricing was not available to fund managers during the period of observation.

Consistent with the predictions from the conceptual framework, we observe that negative returns prompt larger outflows from funds that cannot swing than from their swing pricing counterparts. This observation holds during normal market conditions. Yet during the 2013 U.S. taper tantrum, a period of sharp declines in bond prices, funds appear to have been equally exposed to investor redemptions regardless of whether they applied swing pricing. Even so, swing pricing funds generated higher returns during this episode. This tallies with the predicted anti-dilution effect of swing pricing, which is based on the redistribution of liquidation costs from remaining investors to those deciding to redeem their shares.

We conclude that current swing pricing rules, which tend to apply only a modest swing factor if outflows exceed a certain threshold, fail to offset investor first-mover advantages in stressed markets.

Regulatory responses to address vulnerabilities in the mutual fund industry can thus be enhanced by allowing funds more flexibility in setting swing factors, provided that sound governance policies ensure a transparent and fair treatment of investors. Specifically, swing factors should be an increasing function of liquidity costs and redemptions to offset investor first-mover advantages during episodes of stress. For large funds or groups of funds with common exposures, topping up swing factors to account for the adverse price impact of asset liquidations could further enhance the effectiveness of this tool for macroprudential purposes.

Swing pricing is likely to be most effective if combined with other tools to address first-mover advantages. While swing pricing seeks to mitigate such advantages by reducing the dilution of the NAV, countercyclical liquidity requirements and liquidity stress testing could lower funds' liquidity mismatch and thereby help to further support fund resilience. That said, more research is needed to assess the underlying risks in this industry and the interaction of liquidity-management tools in order to inform the design of a macroprudential framework for the fund-management industry.

References

Abadie, A., and G. W. Imbens. 2011. "Bias-Corrected Matching Estimators for Average Treatment Effects." *Journal of Business and Economic Statistics* 29 (1): 1–11.

Aramonte, S., C. Scotti, and I. Zer. 2020. "Measuring the Liquidity Profile of Mutual Funds." *International Journal of Central Banking* 16 (5): 143–78.

Arellano, M., and S. Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies* 58 (2): 277–97.

Association of the Luxembourg Fund Industry (ALFI). 2015. "Swing Pricing Update 2015." Association of the Luxembourg Fund Industry.

Bank for International Settlements. 2013. "Markets Precipitate Tightening." *BIS Quarterly Review* (September): 1–11.

Bao, J., M. O'Hara, and A. Zhou. 2018. "The Volcker Rule and Market-Making in Times of Stress." *Journal of Financial Economic*s 130 (1): 95–113.

Ben-David, I., J. Li, A. Rossi, and Y. Song. 2022. "What Do Mutual Fund Investors Really Care About?" *Review of Financial Studies* 35 (4): 1723–74.

Ben-Rephael, A. 2017. "Flight-to-Liquidity, Market Uncertainty, and the Actions of Mutual Fund Investors." *Journal of Financial Intermediation* 31 (July): 30–44.

Berk, J., and R. Green. 2004. "Mutual Fund Flows and Performance in Rational Markets." *Journal of Political Economy* 112 (6): 1269–95.

Bessembinder, H., S. Jacobsen, W. Maxwell, and K. Venkataraman. 2018. "Capital Commitment and Illiquidity in Corporate Bonds." *Journal of Finance* 73 (4): 1615–61.

BlackRock. 2016. "Swing Pricing: The Dilution Effects of Investor Trading Activity—Luxembourg Domiciled Funds."

Capponi, A., P. Glasserman, and M. Weber. 2020. "Swing Pricing for Mutual Funds: Breaking the Feedback Loop between Fire Sales and Fund Runs." *Management Science* 66 (8): 3295–3798.

Chen, Q., I. Goldstein, and W. Jiang. 2010. "Payoff Complementarities and Financial Fragility: Evidence from Mutual Fund Outflows." *Journal of Financial Economics* 97 (2): 239–62.

Chernenko, S., and A. Sunderam. 2016. "Liquidity Transformation in Asset Management: Evidence from the Cash Holdings of Mutual Funds." NBER Working Paper No. 22391.

———. 2020. "Do Fire Sales Create Externalities?" *Journal of Financial Economics* 135 (3): 602–28.

Chevalier, J., and G. Ellison. 1997. "Risk Taking by Mutual Funds as a Response to Incentives." *Journal of Political Economy* 105 (6): 1167–1200.

Cominetta, M., C. Lambert, A. Levels, A. Rydén, and C. Weistroffer. 2018. "Macroprudential Liquidity Tools for Investment Funds — A Preliminary Discussion." *Macroprudential Bulletin* (European Central Bank) (October 2): Article 6.

Dannhauser, C., and S. Hoseinzade. 2022. "The Unintended Consequences of Corporate Bond ETFs: Evidence from the Taper Tantrum." *Review of Financial Studies* 35 (1): 51–90.

Edelen, R. M. 1999. "Investor Flows and the Assessed Performance of Open-End Mutual Funds." *Journal of Financial Economics* 53 (3): 439–66.

Falato, A., I. Goldstein, and A. Hortaçsu. 2021. "Financial Fragility in the COVID-19 Crisis: The Case of Investment Funds in Corporate Bond Markets." *Journal of Monetary Economics* 123 (October): 35–52.

Ferreira, M. A., A. Keswani, A. F. Miguel, and S. B. Ramos. 2012. "The Flow-Performance Relationship around the World." *Journal of Banking and Finance* 36 (6): 1759–80.

Financial Stability Board. 2020. "Holistic Review of the March Market Turmoil."

Franzoni, F., and M. C. Schmalz. 2017. "Fund Flows and Market States." *Review of Financial Studies* 30 (8): 2621–73.

Friewald, N., R. Jankowitsch, and M. G. Subrahmanyam. 2012. "Illiquidity or Credit Deterioration: A Study of Liquidity in the US Corporate Bond Market during Financial Crises." *Journal of Financial Economics* 105 (1): 18–36.

Fulkerson, J. A., B. D. Jordan, and T. B. Riley. 2013. "Return Chasing in Bond Funds." *Journal of Fixed Income* 22 (4): 90–103.

Goldstein, I., H. Jiang, and D. T. Ng. 2017. "Investor Flows and Fragility in Corporate Bond Funds." *Journal of Financial Economics* 126 (3): 592–613.

Goldstein, M., and E. Hotchkiss. 2020. "Providing Liquidity in an Illiquid Market: Dealer Behavior in US Corporate Bonds." *Journal of Financial Economics* 135 (1): 16–40.

Greene, J. T., and C. W. Hodges. 2002. "The Dilution Impact of Daily Fund Flows on Open-End Mutual Funds." *Journal of Financial Economics* 65 (1): 131–58.

Hu, G. X., J. Pan, and J. Wang. 2013. "Noise as Information for Illiquidity." *Journal of Finance* 68 (6): 2341–82.

International Organization of Securities Commissions (IOSCO). 2015. "Liquidity Management Tools in Collective Investment Schemes: Results from an IOSCO Committee 5 Survey to Members."

Jiang, H., D. Li, and A. Wang. 2021. "Dynamic Liquidity Management by Corporate Bond Mutual Funds." *Journal of Financial and Quantitative Analysis* 56 (5): 1622–52.

Jin, D., M. Kacperczyk, B. Kahraman, and F. Suntheim. 2022. "Swing Pricing and Fragility in Open-End Mutual Funds." *Review of Financial Studies* 35 (1): 1–50.

Kacperczyk, M., S. van Nieuwerburgh, and L. Veldkamp. 2014. "Time-Varying Fund Manager Skill." *Journal of Finance* 69 (4): 1455–84.

Kargar, M., B. Lester, D. Lindsay, S. Liu, P. O. Weill, and D. Zuiga. 2021. "Corporate Bond Liquidity during the COVID-19 Crisis." *Review of Financial Studies* 34 (11): 5352–5401.

Kashyap, A., D. Kohn, and D. Wessel. 2021. "What Is Swing Pricing?" Up Front blog, Brookings Institution, August 3.

Lewrick, U., and J. Schanz. 2017. "Liquidity Risk in Markets with Trading Frictions: What Can Swing Pricing Achieve?" BIS Working Paper No. 663.

Ma, Y., K. Xiao, and Y. Zeng. 2020. "Mutual Fund Liquidity Transformation and Reverse Flight to Liquidity." Jacobs Levy Equity Management Center for Quantitative Financial Research Paper.

Malik, S., and P. Lindner. 2017. "On Swing Pricing and Systemic Risk Mitigation." IMF Working Paper No. 17/159.

Morris, S., and H. S. Shin. 2003. "Global Games: Theory and Applications." In *Advances in Economics and Econometrics*, ed.

M. Dewatripont, L. Hansen, and S. Turnovsky, 56–114 (chapter 3). Cambridge University Press. The Eighth World Congress.

O'Hara, M., and X. Zhou. 2021. "Anatomy of a Liquidity Crisis: Corporate Bonds in the COVID-19 Crisis." *Journal of Financial Economics* 142 (1): 46–68.

Sirri, E. R., and P. Tufano. 1998. "Costly Search and Mutual Fund Flows." *Journal of Finance* 53 (5): 1589–1622.

Spiegel, M., and H. Zhang. 2013. "Mutual Fund Risk and Market Share-Adjusted Fund Flows." *Journal of Financial Economics* 108 (2): 506–28.

Vivar, L. M., M. Wedow, and C. Weistroffer. 2020. "Burned by Leverage? Flows and Fragility in Bond Mutual Funds." ECB Working Paper No. 2413.

A Pitfall of Cautiousness in Monetary Policy*

Stéphane Dupraz, Sophie Guilloux-Nefussi,
and Adrian Penalver
Banque de France

Central banks are often reluctant to take immediate or forceful actions in the face of new information on the economic outlook. To rationalize this cautious approach, Brainard's *attenuation principle* is often invoked: when a policymaker is unsure of the effects of his policies, he should react less than he would under certainty. We show that the Brainard principle, while a wise recommendation for policymaking in general, runs into a pitfall when it is applied to a central bank setting monetary policy. For a central bank, concerns about uncertainty create a cautiousness bias: acting less is justified when taking as given the private sector's expectations of inflation, but acting less shifts these inflation expectations away from the central bank's inflation target. In response to the de-anchoring of expectations, the central bank can easily end up acting as much as it is initially reluctant to do, but without succeeding in putting inflation back on target. This pattern is a feature of policy under discretion: the central bank would often be better off tying its hands and not respond to its concerns about uncertainty.

JEL Codes: E31, E52, E58.

1. Introduction

Central banks must set monetary policy under substantial uncertainty on the economic outlook, as well as the effects of their own policies. Faced with this uncertainty, they often react by attenuating their policy response, or by changing it only gradually. To justify

*We thank Gadi Barlevy, Mario Pietrunti, Ricardo Reis, and Ulf Söderström for useful discussions, and participants at the CEPR-Bank of Finland conference, Sciences Po, CEBRA conference, Banca d'Italia, and Banco Central de Chile for valuable comments. The views expressed herein are those of the authors and do not necessarily reflect those of the Banque de France or the Eurosystem.

this cautious approach to policymaking, they often refer to Brainard (1967), who formally derived what came to be known as the Brainard principle: although a policymaker who is uncertain of the economic outlook should act as if his best expectation were a sure outcome (Theil 1957), a policymaker who is uncertain of the effects of his own policies should act less than he would under certainty.

The logic of Brainard's attenuation principle is not limited to monetary policy, but it became well known by central bankers in the 1990s thanks to Alan Blinder's book on his experience as a central banker (Blinder 1999).[1] Blinder himself declared that the Brainard principle "was never far from [his] mind when [he] occupied the Vice Chairman's office at the Federal Reserve." More recently, and in the context of an increased reliance of central banks on unconventional policies, Powell (2018) nicely summed up the Brainard logic through the following formula: "When unsure of the potency of a medicine, start with a somewhat smaller dose." On the other side of the Atlantic, in March 2019 Mario Draghi (2019) explained the decision of the European Central Bank (ECB) Governing Council in the following terms: "You just do what you think is right and you temper [with] a consideration [that] there is uncertainty. In other words, in a dark room you move with tiny steps." Bernanke (2007) and Carney (2017) make similar references to the Brainard principle.[2] Other influential policymakers, such as Williams (2013) and Praet (2018), provide more extended analysis of the Brainard principle.[3]

The attention central bankers declare paying to the Brainard principle suggests it affects their monetary policy decisions, and can

[1]Blinder was not only instrumental in popularizing Brainard's principle but also in giving it its current interpretation of a rationale for "doing less." Blinder himself refers to the Brainard principle as the *conservatism principle*. We follow Reinhart (2003) in using the more neutral terminology of *attenuation principle*.

[2]See also Villeroy de Galhau (2018): "In the face of uncertainty, one often hears reference to the celebrated Brainard "conservatism principle" [...] But we should go beyond a static view of Brainard's principle (which focuses on one single small step): a dynamic view would include the time dimension and consider how to manage and communicate a sequence of incremental steps."

[3]"In these conditions, the Brainard rationale for gradualism applies with great force: do it as carefully and prudently as possible, at least when you have good reason to believe that the degree of uncertainty as to the direction and size of market reactions is atypically large. Present times, where policy is defined by a multiplicity of instruments that interact in ways that are very imperfectly known, are characterized by an abnormal amount of uncertainty"(Praet 2018).

therefore contribute to explaining the dynamics of inflation. In recent years, however, the issue has received little attention in the academic literature. Motivated by the prolonged undershooting of the inflation target in the euro area during much of the past decade, we make a new contribution to the theory of monetary policy under uncertainty.

We show that the Brainard principle, while a wise recommendation for policymaking in general, runs into a pitfall when it is applied to a central bank setting monetary policy. For concreteness, we focus on interest rate policies. When a disinflationary shock hits, the central bank can push inflation back up by cutting interest rates. The Brainard principle would recommend that, if the central bank is uncertain of the precise effect of an interest rate cut on inflation, it should cut interest rates by less, even if this means letting inflation fall somewhat below target. This recommendation, however, abstracts from the fact that inflation also depends on the private sector's expectations of inflation, a dimension that Brainard's original setup does not incorporate. The central bank takes these inflation expectations as given when it acts under discretion, but if the private sector foresees that the central bank will attenuate its policy response, it forms lower inflation expectations. This pushes inflation further down, and forces the central bank to decrease rates further. The central bank easily ends up decreasing rates by as much as it is initially reluctant to do, but with an inflation rate further below target than if it had not been concerned about uncertainty.

We give the name *cautiousness bias* to this perverse incentive that turns the central bank's concerns over uncertainty against its own interests. The terminology is in direct reference to the inflation bias expounded by Kydland and Prescott (1977) and Barro and Gordon (1983a, 1983b). Like the inflation bias, the cautiousness bias is a feature of policy under discretion: it arises because the central bank fails to internalize the effect of its policy on inflation expectations. Contrary to the inflation bias, however, it does not arise from a desire by the central bank to set output above its natural level. It does not even require the central bank to care about stabilizing output, and applies equally to a central bank that has a single mandate to stabilize inflation only.

Our analysis of the cautiousness bias is motivated by the inflation dynamics of the euro area in recent years. Over much of the past decade, ECB monetary policy decisions oscillated between a cautious and gradual approach in the face of uncertainty (as exemplified by Draghi's recommendation to "move with tiny steps in a dark room," cited above) and bold decisive actions when inflation expectations started to risk disanchoring (such as the decision to start quantitative easing in January 2015). This dual strategy seems in line with the discussion of the Brainard principle given by Peter Praet in 2018. Praet (2018) argues that "a case for gradualism can be made in the context of the uncertainty inherent in economic data, models and parameters, notably in times of unconventional monetary policy," but that "a more aggressive monetary policy response, however, is warranted when there is clear evidence of heightened risks to price stability." Our analysis does not object to this distinction but warns against taking the risks to price stability as exogenous: the disanchoring of inflation expectations that calls for an aggressive policy response can be precisely caused by the earlier desire to attenuate the policy response.

To show the robustness of the cautiousness bias, we study it under various specifications of the Phillips-curve relationship between output and inflation. In Section 2, we start by explaining its logic with the New Classical Phillips curve. We show that in response to shocks foreseen by the private sector, a cautious central bank ends up moving real rates by exactly as much as a central bank that disregards concerns over uncertainty would. However, despite ending up moving real rates by the same amount as a central bank that disregards concerns over uncertainty (which is also the optimal policy under commitment), a cautious central banker suffers greater departures of inflation from its target. In the spirit of Rogoff (1985)'s solution to the inflation bias, we show that society would be better off appointing a central banker who discounts concerns over uncertainty relative to society, even if this means responding to unforeseen shocks too aggressively.

Although the case of the New Classical Phillips curve provides a simple exposition of the cautiousness bias, its absence of dynamics prevents an analysis of the ongoing interplay between interest rate decisions and the response of inflation expectations. In Section 3, we study the dynamics induced by the cautiousness bias under the sticky-information Phillips curve of Mankiw and Reis (2002). With the sticky-information Phillips curve, the private sector only gradually incorporates new information into its inflation expectations. As a result, when a negative shock hits inflation expectations move little at first, and the central bank is able to attenuate the decrease in interest rates. But as the private sector gradually realizes the resulting below-target inflation, inflation is pushed down further, forcing the central back to decrease rates further, ultimately by as much as if it had not been willing to attenuate policy.

In Section 4 we show that the cautiousness bias does not depend on the sluggish adjustment of expectations in the New Classical Phillips curve and sticky-information Phillips curve. It applies equally to the forward-looking New Keynesian Phillips curve (NKPC), which remains the most commonly used Phillips curve in economic modeling. The timing in the manifestation of the bias is however different in this case, due to the front-loaded dynamics the NKPC is known to generate (Ball 1994; Mankiw and Reis 2002). In response to a persistent fall in the natural rate, agents immediately expect that the central bank will let inflation fall below target in the future. As a result, the central bank is forced to decrease interest rates more, as early as on impact. In this, the dynamics of the cautiousness bias under the NKPC resembles the one under the New Classical Phillips curve.

The cautiousness bias we focus on in most of the paper concerns the response of inflation to the underlying shocks. Although the resulting undershooting or overshooting of the inflation target can be very persistent in the face of very persistent shocks, it does not create an incentive for a discretionary central bank to let *average* inflation depart form the inflation target π^*, in contrast to the inflation bias of Kydland and Prescott. In Section 5, we show that this is only due to an implicit assumption of the frameworks used in previous sections. By generalizing the setup, we show that the conflict between the desire to stabilize inflation and the desire to minimize inflation uncertainty can also lead to an average bias, just as the conflict between the desire to stabilize inflation and the desire to stabilize output can lead to both an inflation bias and a stabilization bias (Svensson 1997).

For concreteness we analyze the cautiousness bias in the context of conventional interest rate policies, but its logic applies equally to unconventional policies such as forward guidance and balance sheet policies—at least when these are intended as alternative ways to stimulate aggregate demand. What is key to the cautiousness bias is not the way monetary policy affects aggregate demand, but the way aggregate demand affects inflation through the Phillips curve. Because unconventional policies are precisely the ones whose effects are likely to be the most uncertain (see, e.g., Williams 2013), the importance for a central bank to be aware of a bias toward excessive caution is all the more important when the effective lower bound (ELB) on nominal interest rates only leaves unconventional policies available.

The cautiousness bias has another implication for unconventional policies. Although in our framework *real* interest rates never move more than if the central bank had not tried to attenuate policy, *nominal* interest rates can. In the presence of the ELB, this implies that a central bank can find itself up against the ELB and forced to turn to unconventional policies even though it would not have, had it not tried to attenuate policy.

A number of papers have considered the implications of model uncertainty for the conduct of monetary policy. Clarida, Gali, and Gertler (1999); Estrella and Mishkin (1999); Svensson (1999); Sack (2000); Sack and Wieland (2000); Rudebusch (2001), and more recently Williams (2013), recover Brainard's recommendation for policy attenuation in the context of monetary policy. Subsequent literature has emphasized situations in which Brainard's attenuation principle is overturned and uncertainty calls instead for a more aggressive response.[4] Söderström (2002), Kimura and Kirozumi (2007), and Ferrero, Pietrunti, and Tiseno (2019) consider such situations while still modeling the central bank's uncertainty in a Bayesian way, as in Brainard's original setup (and ours). Söderström (2002) shows policy aggressiveness can be called for when uncertainty bears on the persistence of inflation, in a model with adaptive expectations. Kimura and Kirozumi (2007) show it is appropriate when uncertainty bears on the fraction of firms that form expectations in a rule-of-thumb, adaptive, fashion. Closer to this paper, Ferrero, Pietrunti, and Tiseno (2019) show that uncertainty about the slope of the New Keynesian Phillips curve can lead the central bank to move nominal interest rates by more than under certainty in response to cost-push shocks, if shocks are persistent enough. We

[4]Brainard (1967)'s original paper already contains situations in which uncertainty calls for more aggressive policy, as we discuss in Section 2. See also Chow (1973), Craine (1979), and Walsh (2003).

interpret the result through the lens of the cautiousness bias: the optimal discretionary policy is to attenuate the policy response for given inflation expectations, but the adverse reaction of inflation expectations forces the central bank to ultimately act more. Overall, our contribution is to show that for a central bank, following the Brainard principle ends up generating excessive deviations of inflation from target simply because agents in the economy are forward looking and understand how such cautiousness affects inflation dynamics.

Other papers consider the consequences of modeling the central bank's uncertainty through the minmax approach of robust control instead of the Bayesian approach. They usually find that uncertainty calls for more aggressive policy, in opposition to the Brainard principle. Giannoni (2002) finds that the fear that the worst will happen to output and inflation if the central bank does not track the natural rate provides an incentive to track it more closely—i.e., to move interest rates more aggressively (see also Stock 1999; Tetlow and von zur Muehlen 2001; Onatski and Stock 2002; Söderström and Leitemo 2008). Sargent (1999) finds that uncertainty about the persistence of shocks calls for a more aggressive response of monetary policy. Barlevy (2011) argues that what is conducive to more aggressive policy under robust control is less the minmax approach per se than its application to specific situations.[5] He gives examples where the minmax approach calls for policy attenuation, for the same reason as under the Bayesian approach, and shows that uncertainty on the persistence of shocks calls for aggressive policy in the Bayesian setup as well.[6]

Other arguments for attenuation or gradualism have been put forward that do not rely on the presence of uncertainty.[7] Woodford (2003c) shows that the optimal, history-dependent, monetary policy

[5]See also Onatski and Williams (2003) and Tillmann (2009).

[6]Using as well the minmax approach of robust control, Woodford (2010) and Adam and Woodford (2012) consider the design of monetary policy rules that are robust to the possibility that the private sector forms expectations using a wrong model, even though the central bank is itself sure of the model of the economy.

[7]Still other arguments have been put forward to explain gradualism positively, without defending it normatively. For instance, Riboni and Ruge-Murcia (2010) argue that gradualism in monetary policy is partly due to the consensus-building approach taken by many monetary policy committees, a decisionmaking procedure that favors the status quo. See also Favaretto and Masciandaro (2016). Spiegler (2021) shows that when the private sector's subjective causal model

under commitment features inertia and can be approximated by a discretionary central bank that puts a cost on abrupt changes in interest rates.[8]

A third argument for gradualism is based on concerns about the stability of the financial system. As argued by Cukierman (1991), interest rate smoothing can be desirable because it mitigates sudden changes in banks' short-term funding costs or long-term asset returns, and therefore in banks' profits and balance sheets. Interestingly, in a recent paper Stein and Sunderam (2018) show that this distinct motive for gradualism can also lead to a time-inconsistency problem: If the central bank dislikes volatile long-term rates for financial stability reasons, it has an incentive to track the natural interest rate only gradually, in order not to reveal information on long-term natural rates that would make long-term rates react too abruptly. But this is taking markets' expectations as given: in equilibrium markets understand that the central bank is moving gradually and adjust their expectations of long-term natural rates accordingly, partly undoing the central bank's efforts. We show that time inconsistency is equally at play when gradualism is driven by uncertainty concerns. The time-inconsistency problem is different between the two models however: in our model the cautiousness bias arises from a failure to internalize inflation expectations, while in Stein and Sunderam's model time inconsistency arises from a failure to internalize expectations of future interest rates—there are no inflation expectations in their model.[9]

---

reverses the direction of causality between inflation and output, the central bank can react by attenuating its policy response relative to when the private sector has rational expectations. Other costs of delaying policy action have also been put forward. Acharya et al. (2019) show that under skill-loss hysteresis, not acting promptly can push the economy into a recession from which it can then only slowly recover, if at all.

[8]Since Woodford's argument for gradualism does not rely on concerns over uncertainty, however, it is not a rationale for attenuating or delaying the policy response more when uncertainty is higher. In particular, there is no rationale for being more reluctant to act when the only instruments available are unconventional instruments with more uncertain effects. As far as Odyssean forward guidance is concerned, it is precisely Woodford's argument in favor of inertia in interest rates that makes committing to keeping rates lower for longer a superior strategy (Eggertsson and Woodford 2003).

[9]As a consequence, the cautiousness bias is at play in our model even though we make the New Keynesian assumption that aggregate demand and inflation

2. A Simple Model of the Cautiousness Bias

In this section, we expose the cautiousness bias in a simple model where the supply side is captured by the New Classical Phillips curve (Lucas 1972). Using the New Classical Phillips curve has two advantages. First, it is the Phillips curve for which the bias appears most transparently. Second, it follows the classic accounts of the inflation bias by Kydland and Prescott (1977) and Barro and Gordon (1983a, 1983b).

2.1 The Problem of the Central Bank

The problem of the central bank is to pick an allocation for inflation π_t, the output gap x_t, and the nominal interest rate i_t that best fits its objective, subject to the constraints imposed by the behavior of the private sector. These constraints are captured by a simple two-equation model. The aggregate-demand side of the economy is represented by the Euler equation:

$$x_t = -\sigma(i_t - E_t(\pi_{t+1})) + E_t(x_{t+1}) + v_t, \tag{1}$$

where σ is the intertemporal elasticity of substitution, and v_t is a possibly autocorrelated exogenous shock with mean zero, observable at period t. The shock v_t captures variations in natural output y_t^n or, equivalently, in the natural rate of interest r_t^n. Specifically, v_t is the function $v_t = -(y_t^n - E_t(y_{t+1}^n))$ of natural output, and connects to the natural rate through $v_t = \sigma r_t^n$. Appendix A derives and discusses the connection between these alternative representations of the fundamental shocks to the Euler equation.

The aggregate supply side of the economy is captured by the New Classical Phillips curve:

$$\pi_t = \kappa x_t + E_{t-1}(\pi_t), \tag{2}$$

where κ is the slope of the Phillips curve. The private sector's past expectations of present inflation, formed at $t-1$, shift the Phillips

depend on the long-term—not short-term—interest rate. In Stein and Sunderam's model, for the time-inconsistency problem to be at play it is necessary that the central bank's interest rate target bears on the short-term rate only.

curve. The New Classical Phillips curve can be derived for instance under the assumption that a fraction of firms set their prices at t with outdated information from $t-1$ (Woodford 2003b; Mankiw and Reis 2010).

As in the literature on the inflation bias which distinguishes between anticipated inflation and surprise inflation (e.g., Barro and Gordon 1983a, 1983b), we allow for the shock v_t to be partially anticipated by the firms that set their prices with outdated $t-1$ information, by letting v_t be partially forecastable with $t-1$ information. Accordingly, we refer to $E_{t-1}(v_t)$ as the *foreseen shocks* and to $v_t - E_{t-1}(v_t)$ as the *unforeseen shocks.*[10]

Crucially, the central bank faces parameter uncertainty. It is uncertain about the value of the structural parameter σ, and entertains several possible values for it. Like Brainard (1967), we follow Savage (1954) in modeling parameter uncertainty in a Bayesian way. The central bank assigns probabilities to every possible value of σ and treats it as a random variable. We note $\bar{\sigma}$ and V_σ the mean and variance of the central bank's subjective beliefs over σ. We assume that V_σ is constant over time. We assume the central bank is certain of the value of κ. We do so because assuming uncertainty bears only on σ is the case most favorable to Brainard's attenuation principle, as will become clear below.

Although the central bank is uncertain of the model of the economy, we assume that the models it entertains are not too far from the actual one, in the spirit of rational expectations. Specifically, we assume that the true value of σ is $\bar{\sigma}$, the mean of the values considered by the central bank. The true dynamics of the economy are therefore given by the Euler equation (1) and Phillips curve (2) with $\sigma = \bar{\sigma}$. Note that we implicitly assume that the private sector is not subject to parameter uncertainty, since we assume that the Euler equation (1) and Phillips curve (2) hold, both of which are derived under the assumption of no parameter uncertainty. As a consequence, the central bank and the private sector have different information sets at t. To avoid any confusion, we denote by $E_t^*(.)$ the expectations of the central bank at t, which are formed without

[10]We take *"anticipated," "foreseen,"* and *"expected"* as synonyms, but reserve *"foreseen"* to the private sector's anticipations of the exogenous shocks and *"expected"* to the private sector's anticipations of inflation.

knowing σ. We reserve the notation $E_t(.)$ for the expectations of the private sector, which knows that $\sigma = \bar{\sigma}$. We assume that the private sector's expectations are part of the central bank's information set. This is meant to capture the fact that central banks have access to—and heavily monitor—measures of the private sector's inflation expectations before taking monetary policy decisions, such as market-based expectations or surveys of professional forecasters.[11]

We assume that the mandate of the central bank is to stabilize inflation only. Its objective is to set inflation π_t to a target π^* at all periods.[12] It has the quadratic loss function:

$$\mathcal{L}_\infty = E_t^* \left(\sum_{k=0}^{\infty} \beta^k (\pi_{t+k} - \pi^*)^2 \right). \tag{3}$$

The assumption of a single inflation mandate is not necessary for our results. Appendix D shows that they hold equally well in the more general case in which the central bank has a dual objective to stabilize both inflation and the output gap. We focus on the case of a single mandate in the body of the paper for two reasons. First, it emphasizes that the cautiousness bias does not arise from a desire to stabilize output at the expense of stabilizing inflation, unlike the

[11]Note that since the private sector's expectations depend on the parameter σ, the central bank could in theory solve for the dependence of expectations on σ and infer the value of σ from expectations. Such an inference is possible in our model because of the simplicity of its stylized two-equation setup and the simplicity of its information structure. The mapping between σ and expectations could be made arbitrary noisy by adding noise to the model, making the inference arbitrarily uninformative. Alternatively, we could assume that the private sector faces the same model uncertainty as the central bank so that the central bank has nothing to learn from the private sector, but at the cost of less standard forms for the Euler equation and Phillips curve. Since the issue is peripheral to our focus, we simply assume away this inference from endogenous signals, as is common in the literature on incomplete perfect knowledge (e.g., Woodford 2003a or Angeletos and La'O 2010).

[12]With the New Classical Phillips curve (2) (and other information-based Phillips curves), the loss function that can be microfounded from the costs of relative price dispersion contains unexpected inflation, not inflation. However, since in practice the mandate of central banks bears on inflation, we assume so here. We consider any arbitrary inflation target (not necessarily zero) for the same reason. Considering other costs of inflation could justify caring about expected inflation (and various values for the inflation target) even when the Phillips curve is based on information frictions.

inflation bias. Therefore, it applies equally to central banks with a single primary mandate, like the ECB. In the appendix, we allow for the central bank to be willing to set output above potential and therefore be subject to the inflation bias, and show that the cautiousness bias and inflation bias arise from distinct perverse incentives. Second, the assumption of a single objective corresponds to the original framework of Brainard (1967).

2.2 Reductio ad Brainard

We now show that this simple monetary model exactly fits into the framework considered by Brainard (1967), up to one key difference: the presence of the expectations of the private sector. In its canonical form, Brainard's model considers a policymaker who seeks to set a single variable on a target through the use of a single instrument. In our case, the single objective is inflation. We pick the single instrument of the central bank to be the real interest rate $r_t \equiv i_t - E_t(\pi_{t+1})$.[13]

By taking expectations at $t-1$ of the New Classical Phillips curve (2), it must be that the expected output gap next period is zero, $E_t(x_{t+1}) = 0$. Plugging in the expression for the output gap from the Euler equation (1) into the Phillips curve (2), we get

$$\pi_t = -\phi r_t + \varepsilon_t + E_{t-1}(\pi_t), \tag{4}$$

where we define $\phi \equiv \sigma\kappa$ and $\varepsilon_t \equiv \kappa v_t$. We denote $\bar{\phi} = \kappa\bar{\sigma}$ the mean of ϕ and $V_\phi = \kappa^2 V_\sigma$ its variance.

Since the relationship (4) only contains period-t variables, the objective of the central bank reduces to setting the interest rate r_t to minimize the present-period loss:

$$\mathcal{L}_t(\varepsilon_t) = E_t^*(\pi_t - \pi^*)^2, \tag{5}$$

[13]To be sure, in practice the central bank sets a path for the nominal interest rate, but the implementation of the optimal policy is an issue distinct from the choice of the optimal policy, which is the one we consider here. The latter is a path for all three variables i_t, π_t, and x_t, subject to the constraints imposed by the Euler equation (1) and Phillips curve (2). Parameterizing the equilibrium through the three variables r_t, π_t, and x_t is simply a convenient change of variables, and one that fits into Brainard's framework.

at all periods t and for all realizations of ε_t, subject to constraint (4). Since there is no ambiguity, we drop the time subscripts. The following lemma takes stock and draws the parallel to Brainard's framework.

LEMMA 1 (Reductio ad Brainard). *The program of the central bank is, for any realization of the shock ε, to pick the interest rate r that minimizes*

$$\mathcal{L}(\varepsilon) = E^*((\pi - \pi^*)^2), \tag{6}$$

subject to

$$\pi = -\phi r + \varepsilon + E_{-1}(\pi), \tag{7}$$

where the central bank observes ε and $E_{-1}(\pi)$, and ϕ is a random variable with mean $\bar{\phi}$ and variance V_ϕ.

Up to the expectations of the private sector $E_{-1}(\pi)$, this is exactly the framework considered by Brainard (1967).

The random variable ϕ captures the policymaker's uncertainty on how its own action r affects its objective π—in our case, the central bank's uncertainty over the interest rate channel. As Brainard emphasized, this is the type of uncertainty that can justify policy attenuation. Uncertainty over ε only—in our case, uncertainty over the natural rate—would result in Theil's certainty equivalence: it would leave the optimal policy unchanged, up to replacing ε by its expected value (Theil 1957). Our assumption that the central bank perfectly observes ε abstracts from this irrelevant form of uncertainty.

We have restricted the uncertainty over the interest rate channel to arise from uncertainty over σ, the elasticity of demand to changes in the real interest rate. It could also arise from uncertainty over κ, the elasticity of inflation to changes in demand. We focus on σ because uncertainty over κ would create correlation between $\varepsilon = \kappa v$ and $\phi = \kappa\sigma$. As Brainard notes, such correlation can turn the recommendation for policy attenuation into a recommendation for policy aggressiveness. Our qualification of the Brainard principle is distinct and does not rely on correlated shocks. Therefore, we restrict uncertainty to σ to focus on the standard case of uncorrelated shocks, which is the one most favorable to policy attenuation.

Our main point is that the presence of the private sector's expectations in the Brainard model (7) makes important changes to its policy recommendations. When the outcome of the policy depends on the expectations of the private sector, we need to distinguish between policy under discretion and policy under commitment. The cautiousness bias is a feature of policy under discretion, when the central bank takes the inflation expectations of the private sector as given. We start with this case.

2.3 Brainard's Attenuation Principle

We first show that if we fix the inflation expectations of the private sector, Brainard's attenuation principle holds unchallenged. To take explicit note of the fact that the central bank does not internalize its impact on expectations, we temporarily denote expectations $e(\pi)$ instead of $E_{-1}(\pi)$. For the moment they do not have to bear any resemblance to equilibrium outcomes.

To understand the trade-off at the heart of Brainard's attenuation principle, it is helpful to decompose the mean squared error in its loss function (6) into the square of the distance of average inflation from its target, and the perceived variance of inflation:

$$\mathcal{L}(\varepsilon) = (E^*(\pi) - \pi^*)^2 + Var^*(\pi), \tag{8}$$

$$\text{where } (E^*(\pi) - \pi^*)^2 = (-\bar{\phi}r + \varepsilon + e(\pi) - \pi^*)^2, \tag{9}$$

$$\text{and } Var^*(\pi) = V_\phi r^2. \tag{10}$$

This expression makes apparent the two—possibly conflicting—objectives of the central bank. It wants to bring its expectation of inflation (conditional on ε) to target, and it wants to minimize the (conditional) variance of inflation. Note that through both objectives—including the one of bringing expected inflation on target given the realization of ε—the goal of the central bank is to minimize the overall variance of inflation.[14]

[14] By setting the conditional expectation of inflation on target for every realization of ε, the central bank minimizes the between variance of inflation. By minimizing the conditional variance of inflation, it minimizes the within variance of inflation. The unconditional loss function is equal to $\mathcal{L} = E(\mathcal{L}(\varepsilon)) = (E(\pi) - \pi^*)^2 + Var(E(\pi|\varepsilon)) + E(Var(\pi|\varepsilon))$, where the second term is the between variance of inflation and the third term is the within variance of inflation. They sum to the total variance of inflation by the law of total variance.

Denote r^s the interest rate that the central bank sets when it faces no parameter uncertainty, $V_\phi = 0$. In this case the central bank can focus on minimizing the first term (9) in its loss function, and can fully stabilize inflation on target by setting[15]

$$r^s = \bar{r}^n + \frac{e(\pi) - \pi^*}{\bar{\phi}}, \tag{11}$$

where $\bar{r}^n$ denotes the natural rate in the average model:

$$\bar{r}^n \equiv \frac{\varepsilon}{\bar{\phi}} = \frac{v}{\bar{\sigma}}. \tag{12}$$

According to Equation (11), without concerns for parameter uncertainty, the optimal discretionary policy is to track the natural rate, plus a corrective term if inflation expectations are not on target. In this case, it is by responding fully to variations in the natural rate that the central bank reduces inflation volatility and stabilizes the economy.

When the central bank is uncertain of the impact of its rate decision on inflation, $V_\phi > 0$, the policy rate r affects not only the expected value of inflation (9) but also its variance (10). This new dependence captures the fact that, if the central bank is unsure of the consequences of departing from the steady-state rate $r = 0$, its uncertainty is all the greater the larger the departure away from the steady-state rate. Because the policy rate now affects both terms, there is now a trade-off between reaching the inflation target on average (and minimizing the between variance of inflation) and minimizing the variance of inflation. The Brainard principle answers the question of how the central bank solves this trade-off. It can be obtained by taking the first-order condition of the loss function (8).

[15]The result for r^s can equivalently be written $r^s = \kappa v + e(\pi) - \pi^*$. The expression in the text uses the fact that $\bar{\phi} = \kappa\bar{\sigma}$. Note that this decision-theoretic result is valid regardless of whether the central's bank subjective (average) beliefs $\bar{\phi}$ and $\bar{\sigma}$ correspond to the equilibrium ones. The assumption that the central bank's average model corresponds to the equilibrium one—the standard rational expectations assumption—will only intervene later on when we solve for inflation expectations. The same remark applies to the expression of α in Equation (14) below.

LEMMA 2 (Brainard's Attenuation Principle). *Under discretion, the central bank sets the real interest rate as*

$$r = \alpha r^s, \tag{13}$$

$$\text{where } \alpha \equiv \frac{\bar{\phi}^2}{\bar{\phi}^2 + V_\phi}. \tag{14}$$

The central bank solves the trade-off by choosing a midpoint r between the optimal interest rate policy without parameter uncertainty r_s which minimizes the first term, and the steady-state interest rate $r = 0$ which minimizes the second term. Policy becomes biased toward the steady-state interest rate $r = 0$ because the central bank understands the effects of this policy better. Crucially, because α is less than one, the central bank no longer fully reacts to shocks to the natural rate. Uncertainty over the effects of the policy response calls for attenuating the policy response.

Note that under its optimal discretionary policy the central bank does not expect inflation to be on target. Plugging the chosen policy rate (13) into (7), the central bank expects inflation to be

$$E^*(\pi) = \pi^* + (1-\alpha)(\bar{\phi}\bar{r}^n + e(\pi) - \pi^*) \neq \pi^*. \tag{15}$$

But the central bank is fine with this. It sees it as a cost worth paying to avoid the risks of uncertain policy outcomes.

2.4 The Reaction of Inflation Expectations

The conclusion that the central bank reacts less to shocks is premature, however. The optimal discretionary policy (13) depends on the private sector's expectations of inflation, which are still to be solved for. Crucially, inflation expectations depend on what policy the private sector expects the central bank to implement. If a central bank concerned with parameter uncertainty fights inflation less aggressively, private agents are likely to take it into account in forming inflation expectations.

We solve for the rational expectations of the private sector. Injecting policy (13) into Equation (7), taking expectations E_{-1}, and imposing rational expectations $e(\pi) = E_{-1}(\pi)$ yields

$$E_{-1}(\pi) = \pi^* + \left(\frac{1}{\alpha} - 1\right)\bar{\phi}E_{-1}(\bar{r}^n). \tag{16}$$

When the central bank faces no parameter uncertainty $\alpha = 1$, inflation expectations are on target $E_{-1}(\pi) = \pi^*$, since the private sector rightly anticipates that the central bank will set the policy rate so that inflation is on target under all circumstances. With parameter uncertainty $\alpha < 1$, however, expectations of a natural rate below average leads the private sector to expect below-target inflation. The private sector rightly expects that in this case the cautious central bank will set the real interest rate above the natural rate, creating a negative output gap, and thus below-target inflation.

By plugging the private sector's expectations of inflation in equilibrium (16) into the expression for the policy rate chosen by the central bank (13), we obtain the value of the real interest rate in equilibrium.

Proposition 1 (Brainard Principle Unraveled). *Under the optimal discretionary policy, the real interest rate is in equilibrium:*

$$r = E_{-1}(\bar{r}^n) + \alpha(\bar{r}^n - E_{-1}(\bar{r}^n)). \tag{17}$$

In equilibrium, a central bank with concerns over uncertainty ($\alpha < 1$) attenuates its response only to changes in the natural rate unforeseen by the private sector. To foreseen changes it reacts exactly as much as if it had no concerns about uncertainty.

A cautious central bank ends up reacting just as much to shocks foreseen by the private sector because its reluctance to act pushes inflation to the point at which it is forced to act to the same extent anyway. Assume a shock hits that pushes the natural rate below its average level. Worried by the uncertainty induced if it decreases its policy rate to $\bar{r}^n$, the central bank decides not to fully track the decrease in the natural rate, even if it means letting inflation fall somewhat below target. But if the shock is foreseen by the private sector, this willingness of the central bank to tolerate below-target inflation is, too. Accordingly, the private sector expects lower inflation. Lower inflation expectations put further downward pressure on inflation. In response, the central bank decides to decrease its rate a little more but still in an attenuated manner, which justifies even lower inflation expectations, and so on. Ultimately, inflation expectations fall to the point where they are low enough to convince the central bank to fully match the decrease in the natural rate, as it would have chosen in the absence of concerns for uncertainty.

2.5 The Cautiousness Bias

In reaction to foreseen shocks, the central bank ends up moving its policy rate by as much as if it had no concerns about uncertainty, but inflation ends up further away from target. Formally, the overall departure from target, in response to both unforeseen shocks $\bar{r}^n - E_{-1}(\bar{r}^n)$ and foreseen shocks $E_{-1}(\bar{r}^n)$, is

$$E^*(\pi) - \pi^* = (1-\alpha)\bar{\phi}\left((\bar{r}^n - E_{-1}(\bar{r}^n)) + \frac{1}{\alpha}E_{-1}(\bar{r}^n)\right). \quad (18)$$

In response to unforeseen shocks $\bar{r}^n - E_{-1}(\bar{r}^n)$, inflation ends up away from target, by $(1-\alpha)\bar{\phi}$ percentage points for every percentage-point change in the natural rate. This is the amount of inflation the central bank was willing to tolerate. But in response to foreseen shocks $E_{-1}(\bar{r}^n)$ inflation ends up further away from target, by an additional factor $1/\alpha$. The outcome in terms of stabilizing inflation is worse than if the central bank ignored policy uncertainty. Since the policy rate is forced into territory the central bank was seeking to avoid, the outcome is as bad in terms of avoiding uncertain outcomes. Overall, in response to foreseen shocks the central bank reaches a worse outcome than if it had not sought to act cautiously. Figure 1 provides a graphical illustration similar to the diagrammatic exposition of the inflation bias by Kydland and Prescott (1977).

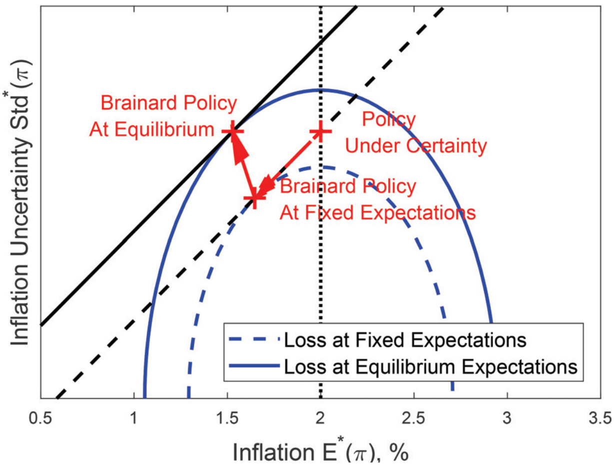
The result that the central bank is behaving against its own interest is a feature of policy under discretion. It would not be if policy were chosen under commitment, because this would allow the central bank to internalize the effect of its policy on expectations. Indeed, Appendix B shows the following result.

PROPOSITION 2 (The Cautiousness Bias). *In the optimal allocation under commitment, the policy rate takes the exact same value as under discretion (17), but inflation departs from target by only*

$$E^*(\pi) - \pi^* = (1-\alpha)\bar{\phi}\left(\bar{r}^n - E_{-1}(\bar{r}^n)\right). \quad (19)$$

Concerns over uncertainty make the discretionary policy depart from the optimal commitment policy.

Figure 1. Consequences of Attenuating Policy when Expectations React



Inflation takes the same value as under discretion in response to unforeseen shocks, but it remains on target in response to foreseen shocks. Under commitment the central bank understands that when shocks are fully foreseen by the private sector, the policy rate can only be equal to the natural rate in equilibrium. It understands that a desire to vary the policy rate by less than the natural rate will only increase inflation expectations up to the point where inflation is enough off target to convince the central bank to vary the policy rate by the full extent of the change in the natural rate. As a consequence, it does not attempt to attenuate the policy response to foreseen shocks and inflation remains on target. It does attenuate the response to unforeseen shocks, however, because these do not risk de-anchoring inflation expectations.

We give the name cautiousness bias to the perverse incentive that turns the central bank's cautiousness into a policy that is no less aggressive, but yields worse stabilization outcomes. Like the inflation bias expounded by Kydland and Prescott (1977) and Barro and Gordon (1983a, 1983b), the cautiousness bias arises because policy chosen under discretion abstracts from the effect of policy on expectations. It differs from the inflation bias, however, in that it does not rely on the desire of the central bank to set output above its natural level. As our assumption of a single inflation mandate highlights, it does not even require the central bank to care about stabilizing output. It arises instead because of the distorted perception of the trade-off between stabilization inflation, and stabilizing the policy rate at values where its effects are better known.[16]

2.6 Guarding Oneself Against Being Cautious

Short of shifting to deciding policy under commitment, what can a central bank—or the society that appoints it—do to guard itself against the cautiousness bias? Rogoff (1985) proposed a solution to realign the incentives of a discretionary policymaker with the preferences of society under commitment: appoint a policymaker whose preferences differ from society's.

We show that society would be better off appointing a central banker who is less cautious than society is. We capture different degrees of cautiousness through different weightings δ of the variance term in the loss function (8):[17]

$$\mathcal{L}(\varepsilon) = (E^*(\pi) - \pi^*)^2 + \delta Var^*(\pi). \tag{20}$$

A discretionary central bank with such preferences still sets policy according to (13), up to a new value for the attenuation coefficient α:

$$\alpha(\delta) = \frac{\bar{\phi}^2}{\bar{\phi}^2 + \delta V_\phi}. \tag{21}$$

[16]Note that the perverse incentive of the cautiousness bias would apply similarly if the central bank's motive for limiting fluctuations in the policy rate—the term in r^2 in its loss function—was driven by something else than concerns over uncertainty—for instance, by concerns over financial stability. Our paper is concerned with Brainard uncertainty, as it is a recurring argument in central bankers' discussions of attenuation and gradualism, but the model can be fruitfully applied to other motives.

[17]The weighting parameter δ allows to encompass several reasons for different degrees of cautiousness. A central banker who perceives less uncertainty on the model parameter ϕ will be akin to one who has a lower δ, since $Var^*(\pi) = V_\phi r^2$. But differences in the degree of cautiousness can reflect pure differences in preferences: the weighting parameter δ can also be seen as capturing different degrees of risk aversion within a class of (squared) mean-variance preferences. Finally, discounting the variance term can be a conscious decision to discount uncertainty concerns in order to counterbalance the cautiousness bias. In this last case, it resembles a form of limited commitment.

We assume that society's true preferences are still captured by the loss function (8), i.e., the loss function (20) with $\delta = 1$, and evaluate the outcome delivered by the various central bankers—different values of δ—according to society's true social preferences.[18] Appendix C shows the following result.

Proposition 3 (Optimal Discounting of Uncertainty Concerns).

- *Unless all shocks are unforeseen by the private sector, it is always desirable to have a central bank that discounts concerns over uncertainty,* $\delta < 1$.
- *The optimal value of* δ *decreases with the proportion of shocks that are foreseen by the private sector.*

A central bank that discounts uncertainty more reacts to shocks more. The benefit is that such a central bank reacts more to foreseen shocks, reducing the cautiousness bias. The cost is that it overreacts to unforeseen shocks. The optimal δ strikes a balance between costs and benefits.

3. The Cautiousness Bias with the Sticky-Information Phillips Curve

We now consider how a more realistic model of the dynamics of inflation affects the workings of the cautiousness bias. While the New Classical Phillips curve is useful to illustrate the logic of the cautiousness bias, its absence of dynamics misses an analysis of the chronology in the policy decisions and their consequences. With the New Classical Phillips curve, the entire dynamics is subsumed into a one-period simultaneous equilibrium: because the private sector anticipates that the central bank will fight deflationary shocks less aggressively at t, inflation expectations are lower at t, and the central

[18]The fact that concerns over uncertainty derive from society's concerns over uncertainty can conflict with our assumption that the private sector faces no parameter uncertainty. As explained in footnote 11, we only make the latter assumption to avoid encumbering the model with a peripheral signal-extraction problem. Alternatively, the social preferences can be interpreted as the preferences of a government which faces the same parameter uncertainty as the central bank.

bank is forced to act at t. To capture the dynamics in a more realistic way, we turn to the sticky-information Phillips curve of Mankiw and Reis (2002). Under the sticky-information Phillips curve, the sequence of events happens sequentially. A negative shock to the natural rate hits; the central bank does not fully track the fall in the natural rate; inflation falls below target; the private sector gradually realizes that inflation is below target and forms lower inflation expectations; lower inflation expectations push inflation down; the central bank is forced to decrease rates further. However, while a cautious central bank can initially attenuate its policy, it still eventually ends up acting as much as if it did not try to attenuate its policy.

3.1 The Problem of the Central Bank

We assume the supply side of the economy is captured by the sticky-information Phillips curve:

$$\pi_t = \kappa x_t + \bar{E}_{t-1}(\pi_t + \zeta \Delta x_t), \tag{22}$$

where ζ is the slope of the short-run aggregate supply (SRAS), $\Delta x_t = x_t - x_{t-1}$ is the growth rate of the output gap, and $\bar{E}_{t-1}$ is notation for the following weighted average of expectations formed at different periods in the infinite past:

$$\bar{E}_{t-1}(\pi_t + \zeta \Delta x_t) = \sum_{j=0}^{\infty} \lambda(1-\lambda)^j E_{t-1-j}(\pi_t + \zeta \Delta x_t). \tag{23}$$

Like the New Classical Phillips curve, the sticky-information Phillips curve models monetary non-neutrality as arising from price setters' imperfect information. Contrary to the New Classical Phillips curve, which assumes all information is incorporated by everyone after one period, the sticky-information Phillips curve assumes that the private sector only gradually learns about the shocks that hit the economy.[19]

[19]Specifically, it can be derived under the assumptions that a fraction λ of firms update their information sets every period, and that the probability of updating information in a given period is independent of how long it has been since the

We only replace the Phillips curve and keep the Euler equation (1) unchanged on the aggregate-demand side. One issue that did not arise in the case of the New Classical Phillips curve is whether to assume that the central bank observes the private sector's current expectation of the output gap tomorrow, $E_t(x_{t+1})$.[20] Assuming one way or the other does not change the results qualitatively. We focus on the case where the central bank does not observe $E_t(x_{t+1})$, as it likely better captures the uncertainty of the central bank on the effect of its policy. Indeed, iterating the Euler equation forward, note that

$$x_t = -\sigma R_t + E_t\left(\sum_{k=0}^{\infty} v_{t+k}\right), \tag{24}$$

where R_t is the long-term real interest rate:

$$R_t \equiv E_t\left(\sum_{k=0}^{\infty} r_{t+k}\right) = r_t + E_t(R_{t+1}). \tag{25}$$

The iterated Euler equation (24) highlights that aggregate demand depends on the effect of the whole sequence of future rates summarized by the long-term real interest rate R_t, i.e., the entire yield curve and not just the short-term real interest rate r_t. In the recursive Euler equation (1), the term $E_t(x_{t+1})$ sums up the effect of the entire yield curve from tomorrow on $E_t(R_{t+1})$ (and of future disturbances) because aggregate demand tomorrow also depends on the entire yield curve from tomorrow on:

$$E_t(x_{t+1}) = -\sigma E_t(R_{t+1}) + E_t\left(\sum_{k=1}^{\infty} v_{t+k}\right). \tag{26}$$

Therefore assuming that the central bank observes $E_t(x_{t+1})$ is akin to assuming that the central bank faces no uncertainty as to the

firm last updated its information. The slope of the sticky-information Phillips curve is a function of the probability λ of renewing one's information set in a given period, $\kappa = \zeta\lambda/(1-\lambda)$.

[20] The issue does not arise with the New Classical Phillips curve because with the New Classical Phillips curve the expected output gap next period is always zero $E_t(x_{t+1}) = 0$.

effect of future short-term real interest rates on current aggregate demand and is only uncertain about the effect of the current short-term real interest rate r_t on current aggregate demand. Since aggregate demand depends on the entire yield curve R_t, assuming that the central bank is equally uncertain about the effect of the entire yield curve R_t is more natural, and we therefore focus on this case. It is handled easily by using the Euler equation in its iterated form (24). Again, the case where the central bank faces uncertainty only on the effect of the short-term interest rate r_t is qualitatively similar.[21]

Plugging in the iterated Euler equation (24) into the sticky-information Phillips curve (22) gives the relationship between the long-term interest rate R_t and inflation π_t at t:

$$\pi_t = -\phi R_t + \kappa E_t \left(\sum_{k=0}^{\infty} v_{t+k} \right) + \bar{E}_{t-1}(\pi_t + \zeta \Delta x_t), \tag{27}$$

where $\phi = \kappa\sigma$. The problem of the central bank under discretion at t is to pick the short-term interest rate $r_t = R_t - E_t(R_{t+1})$ that minimizes the loss $\mathcal{L}_t = E_t^*((\pi_t - \pi^*)^2)$ subject to the constraint (27). Because it acts under discretion, it takes future policies $E_t(R_{t+1})$ as given. As a result, the problem can be phrased equivalently as the one of choosing the long-term interest rate R_t that minimizes the loss $\mathcal{L}_t$.

3.2 The Attenuation Principle in a Dynamic Setup

The central bank still faces the same trade-off as under the New Classical Phillips curve, between bringing its best expectations of

[21] An additional mechanism appears in this case, which tends to make the policy response more front-loaded. In response to a fall in the natural rate, the central bank can expect it will attenuate the decrease in the short-term policy rate tomorrow. As a result, the output gap today becomes more negative and inflation today falls more below target. The central bank is therefore more willing to decrease the short-term interest rate today in order to stabilize inflation today. It is forced into action even earlier on.

inflation on target and minimizing the (conditional) variance of inflation. Its loss function can be written

$$\mathcal{L}_t = (E_t^*(\pi_t) - \pi^*)^2 + Var_t^*(\pi_t), \tag{28}$$

$$\text{where } (E_t^*(\pi_t) - \pi^*)^2 = \left(- \bar{\phi} R_t + \kappa E_t \left(\sum_{k=0}^{\infty} v_{t+k} \right) \right.$$
$$\left. + \bar{E}_{t-1}(\pi_t - \pi^* + \zeta \Delta x_t) \right)^2, \tag{29}$$

$$Var_t^*(\pi_t) = V_\phi R_t^2. \tag{30}$$

The long-term real interest rate that sets inflation on target—the one the central bank would set absent concerns over uncertainty—is

$$R_t^s = R_t^n + \frac{\bar{E}_{t-1}(\pi_t - \pi^* + \zeta \Delta x_t)}{\bar{\phi}}, \tag{31}$$

where we define again the natural rate in the average model $r_t^n = v_t/\bar{\sigma}$, and the long-term natural rate

$$R_t^n \equiv E_t \left(\sum_{k=0}^{\infty} r_{t+k}^n \right). \tag{32}$$

Without concerns over parameter uncertainty, the optimal discretionary policy is to track the natural rate, plus a corrective term if inflation expectations are not on target.[22] In doing so, it has the long-term rate track the long-term natural rate, plus a corrective term if inflation expectations are not on target.

When the central bank is uncertain about the impact of interest rates on inflation, it faces a trade-off between setting inflation on target on average and minimizing the variance of inflation (30). Taking the first-order condition to the loss function (28) gives that the central bank solves this trade-off by setting the policy rate such

[22]This optimal policy tracks the *long-term* natural rate. As a result, if monetary policy is expected not to track the short-term natural rate tomorrow, optimal discretionary policy requires to make the *short-term* rate today depart from the short-term natural rate today in order to align the long-term rate on the long-term natural rate.

that the long-term rate R_t is at a midpoint between the interest rate R_t^s that puts inflation on target on average, and the interest rate $R_t = 0$ that minimizes the variance of inflation:

$$R_t = \alpha R_t^s, \tag{33}$$

where α is still given by (14). Brainard's attenuation principle materializes once again as a bias toward the policy whose effects the central bank understands best: in our case, keeping long-term interest rates toward their steady-state value.

3.3 Acting Tomorrow Out of Not Acting Today

To assess whether—and for how long—the central bank can indeed attenuate its policy response, we need to solve for the private sector's expectations on which the policy rate (33) depends. In order to do so, Appendix E solves for the dynamics induced by this very policy.

Proposition 4 (Dynamics under the Sticky-Information Phillips Curve). *The dynamics of inflation and the output gap (in the average model) are determined by the system:*

$$x_t = \bar{\sigma} R_t^n - \frac{\alpha}{(1-\alpha)\kappa}(\pi_t - \pi^*), \tag{34}$$

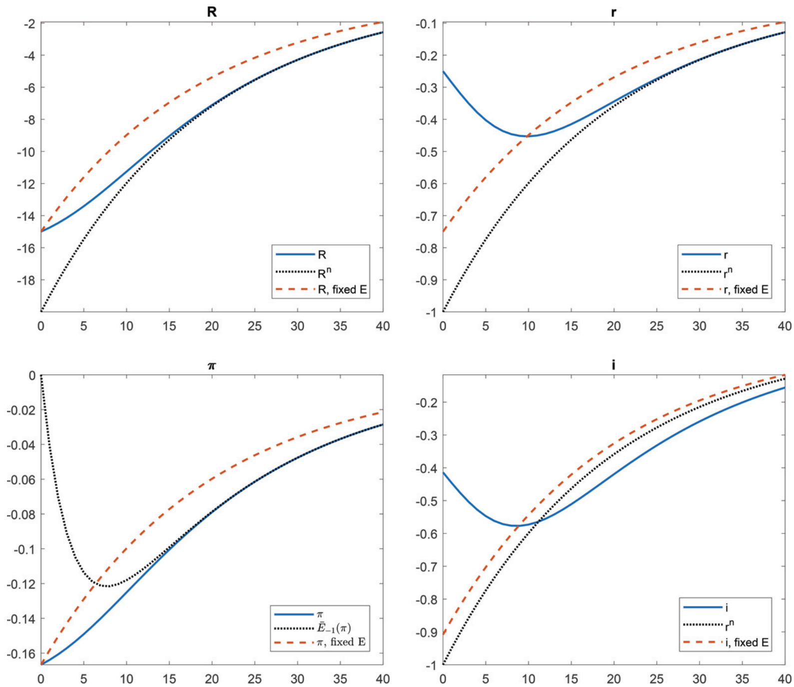
$$\pi_t = \kappa x_t + \bar{E}_{t-1}(\pi_t + \zeta \Delta x_t). \tag{35}$$

We consider the response of the economy (34)–(35) to a persistent fall in the natural rate. We assume the shocks to the natural rate follow an AR(1):

$$r_t^n = \rho r_{t-1}^n + \eta_t, \tag{36}$$

in which case the long-term natural rate $R_t^n = r_t^n/(1-\rho)$ also does. We calibrate the model at a quarterly frequency, as follows. Following Mankiw and Reis (2002), we set the slope of the SRAS to $\zeta = 0.1$, and the frequency of renewing one's information to once a year, $\lambda = 0.25$. This gives a slope of the Phillips curve equal to $\kappa = \zeta\lambda/(1-\lambda) = 0.033$. We set the intertemporal elasticity of substitution to $\bar{\sigma} = 1$. We assume that the uncertainty of the central bank is such that it attenuates its response by a quarter, $\alpha = 0.75$. We set the persistence of the shocks to $\rho = 0.95$.

Figure 2. IRF to a Fall in the Natural Rate under the Sticky-Information Phillips Curve



Note: The figure gives the impulse response functions (IRFs) of the long-term real interest rate R, the short-term real interest rate r, inflation π, and the short-term nominal interest rate i to a 1 percentage point decrease in the real natural rate of interest. The dashed lines give the IRFs in the counterfactual case where expectations of inflation and output gap growth $\bar{E}_{t-1}(\pi_t + \zeta \Delta x_t)$ remain constant at π^*. The horizon is expressed in quarters. The IRFs are plotted under the following calibration. The intertemporal elasticity of substitution is $\bar{\sigma} = 1$; the probability of renewing one's information set in the quarter is $\lambda = 0.25$; the slope of the short-run aggregate supply relationship is $\zeta = 0.1$. It implies a slope of the Phillips curve $\kappa = \zeta \times \lambda/(1-\lambda) \backsimeq 0.033$. The uncertainty V_ϕ is such that the central bank attenuates its action by a quarter, $\alpha = 0.75$. The autoregressive root of the AR(1) shock process is $\rho = 0.95$.

We consider a fall of the natural rate by 1 percentage point on impact. We solve for the impulse response function (IRF) using the method of undetermined coefficients, as detailed in Appendix E. Figure 2 gives the responses of interest rates and inflation. As can be seen through the dotted line of the bottom-left panel, on impact inflation expectations stay close to target because most private agents do not notice the shock. As a result, as shown on the upper-left panel, on impact the central bank is able to decrease the long-term real interest rate R (in plain line) by less than the fall in the natural long-term interest rate R^n (in dotted line). Because the long-term interest rate is above its natural level, inflation falls below target, as can be seen through the plain line of the bottom-left panel. As the private sector gradually realizes the fall in inflation, inflation (in plain line) is gradually pushed down below the path it would have taken had expectations stayed fixed (in dashed line). It forces the central back to decrease rates further. Ultimately, the real long-term interest rate ends up tracking the fall in the natural long-term rate as much as if the central bank had not been willing to attenuate policy and had simply tracked the natural rate from the start. Even as the real rate converges to the natural rate, however, inflation remains below target whereas it would have remained on target if the central bank had not been concerned with uncertainty and had tracked the natural rate from the start.[23]

On Figure 2, although the fall in the real rate never exceeds the fall in the natural rate, because of the fall in inflation expectations the *nominal* rate ends up falling by more that it would have absent concerns over uncertainty. Taking into account the effective lower bound on interest rate policies, a central bank subject to the cautiousness bias can therefore find itself up against the ELB even though it would not have if it had not tried to attenuate policy early on.

4. The Cautiousness Bias with the New Keynesian Phillips Curve

Both the New Classical Phillips curve and the sticky-information Phillips curve make past expectations of present inflation the

[23]Notice that the decrease in the *short-term* real interest rate r is initially even more attenuated than if inflation expectations stayed anchored, as can be seen on the top-right panel. This is because, as private agents anticipate that the central bank will be forced to decrease rates further in the future, initially the central bank faces a lower yield curve. As anticipations of future short-term rates are low, the central bank has less of an incentive to decrease present short-term rates. Ultimately, however, the short-term real interest rate ends up tracking the natural short-term rate as well.

relevant inflation expectations. Does the cautiousness bias depend on this form of sluggish expectations in the Phillips curve? We show it does not: The cautiousness bias applies equally to the forward-looking New Keynesian Phillips curve (NKPC). The dynamics of events under the NKPC differ from those under the sticky-information Phillips curve. Because the NKPC is not based on the assumption that it takes time for agents to incorporate new information, inflation expectations are not sluggish. Monetary policy is not progressively forced into action as inflation expectations progressively adjust. Instead, inflation expectations respond strongly on impact, and monetary policy is forced into action on impact.

4.1 The Problem of the Central Bank

We assume the supply side of the economy is captured by a standard New Keynesian Phillips curve, where inflation expectations enter as present expectations of future inflation:

$$\pi_t = \kappa x_t + \beta E_t(\pi_{t+1}), \tag{37}$$

where $\beta \in (0,1)$ is the discount factor. We keep the Euler equation unchanged, again in its iterated form (24). Because the New Keynesian Phillips curve is slightly non-vertical in the long-run $\beta < 1$, we focus on the case of a zero-inflation target $\pi^* = 0$ in order to abstract from a desire to exploit a long-run trade-off between inflation and output. Plugging the iterated Euler equation (24) into the New Keynesian Phillips curve (37) gives the relationship between the short-term interest rate r_t chosen at t and inflation π_t at t:

$$\pi_t = -\phi\left(r_t + E_t(R_{t+1})\right) + \kappa E_t\left(\sum_{k=0}^{\infty} v_{t+k}\right) + \beta E_t(\pi_{t+1}), \tag{38}$$

where $\phi = \kappa\sigma$. The problem of the central bank under discretion at t is to pick the short-term interest rate r_t that minimizes the loss $\mathcal{L}_t = E_t^*(\pi_t^2)$ subject to this constraint (38). Because it acts under discretion, it takes future policies $E_t(R_{t+1})$ as given.

4.2 The Cautiousness Bias with Forward-Looking Inflation Expectations

The derivation of the discretionary policy is similar to the case of the sticky-information Phillips curve. Appendix F shows that the Brainard principle still applies: the optimal discretionary policy is to attenuate the response of the long-term interest rate to changes in the long-term natural rate by the factor α,

$$R_t = \alpha \left(R_t^n + \frac{\beta E_t(\pi_{t+1})}{\bar{\phi}} \right). \tag{39}$$

Once again, however, the central bank acts less only for given inflation expectations, and acting less shifts inflation expectations adversely. Appendix F solves for inflation expectations and shows that the cautiousness bias applies equally to the New Keynesian Phillips curve.

PROPOSITION 5 (The Cautiousness Bias under the New Keynesian Phillips Curve). *Assume that the natural rate follows an AR(1) process (36). In equilibrium the long-term rate is*

$$R_t = \alpha \left(\frac{1}{1 - \beta(1-\alpha)\rho} \right) R_t^n. \tag{40}$$

While Brainard's attenuation principle leads the central bank to move rates less by a factor $\alpha < 1$, the reaction of inflation expectations forces the central bank to move rates more by a factor $1/(1 - \beta(1-\alpha)\rho) > 1$.

The timing in the manifestation of the cautiousness bias is specific to the New Keynesian Phillips curve, however. A well-known property of the NKPC is that it produces front-loaded dynamics, where shocks have their maximal effect on impact (Ball 1994; Mankiw and Reis 2002). This applies to the dynamics of the cautiousness bias. When a persistent negative shock to the natural rate hits, agents immediately factor in that the central bank will underreact to the fall in the natural rate, letting inflation fall below target. As a result, their present expectations of future inflation—the ones that enter the New Keynesian Phillips curve—immediately fall. In reaction, the central bank is immediately forced to decrease interest rates further in order to counteract the fall in inflation expectations. The central bank is forced into action as early as on impact, and from then on the change to the real interest rate fades away in proportion to the change in the natural rate.

Because the more persistent a shock is, the more inflation expectations react, persistent shocks force the central bank to react the most vigorously. Inflation expectations do not react at all if the shock is fully transitory $\rho = 0$, and react most when ρ tends to one. In our model, the central bank always act less than if it did not have concerns about uncertainty, if acting more is defined in terms of the *real* interest rate. Even for very persistent shocks, the overall multiplier $\alpha/(1-\beta(1-\alpha)\rho)$ always remains below one.[24] Appendix G shows that the same is true when the relevant Euler equation is the recursive Euler equation (1). Under either assumption on the Euler equation, however, for persistent enough shocks, the central bank moves *nominal* interest rates by more than it would have absent concerns about uncertainty, as shown in Appendices F and G.

That the persistence of shocks mitigates the policy attenuation called for by the Brainard principle is emphasized by Ferrero, Pietrunti, and Tiseno (2019). They consider a model where parameter uncertainty applies to the slope of the New Keynesian Phillips curve κ and show that the response of the central bank's nominal interest rate to cost-push shocks can move from attenuated to accentuated if shocks are persistent enough. We interpret their result through the lens of the cautiousness bias: a central bank concerned with Brainard uncertainty always wants to act less, but under discretion, the adverse reaction of the private sector's expectations can force it to act more.

5. Consequences of Generalizing the Least Uncertain Policy

In all the analysis so far, we have implicitly assumed that the interest rate policy whose effects are least uncertain is the steady-state interest rate. As a consequence, the cautiousness bias has not challenged

[24]Since the short-term rate is $r_t = \alpha\left(\frac{1}{1-\beta(1-\alpha)\rho}\right) r_t^n$, the same is true of the short-term interest rate.

the ability of the central bank to set average inflation on target, in contrast to the inflation bias. By generalizing the setup to allow for any policy to be the least uncertain, we show that the conflict between the desire to stabilize inflation and the desire to minimize inflation uncertainty can lead not only to an overreaction bias but also to an average bias, just as the conflict between the desire to stabilize inflation and the desire to stabilize output can lead to both an inflation bias and a stabilization bias (Svensson 1997). We also show that making the least uncertain policy a function of recently implemented interest rates does not change the equilibrium level of real interest rates, provided the private sector observes past policy rates and adjusts its expectations accordingly.

5.1 Allowing for Any Policy to Be the Least Uncertain

We consider again the framework of Section 2, where the economy's supply side is captured by the New Classical Phillips curve (2). In the setup of Section 2, the unconditional average inflation rate ends up equal to the inflation target, $E(\pi) = \pi^*$. This can be seen by taking the unconditional average of the expression for expected inflation (16), and using the fact that the natural rate is at its steady-state value on average, $E(\bar{r}^n) = 0$. While the cautiousness bias makes inflation depart from target by more, departures from target are symmetric above and below the target. Average inflation remains on target.[25]

The absence of an average bias is however only due to an implicit assumption embedded in Equation (7). The rationale for Brainard's attenuation principle is that, if the central bank is uncertain of the effects of its own action on inflation, uncertainty on inflation is all the greater the more it acts. *Acting more,* however, is only defined relative to a reference point. In Brainard's framework, this reference point is the policy whose effects on inflation are best understood $\underline{r}$, in the sense of minimizing the conditional variance of inflation (10). Equation (7) de facto assumes that the policy whose effects are best understood is keeping the interest rate around the steady-state value

[25]Of course, the fact that average inflation is on target also depends on the fact that we shut down the inflation bias by considering a central bank with a single inflation mandate. See Appendix D for the case where both the cautiousness and inflation biases are potentially at play.

of the natural interest rate, $\bar{r}^n = 0$. This is a justifiable assumption. Because the steady-state natural rate is the interest rate that has been most often implemented, it can be argued it is the interest rate on which most experience has been acquired.

But this is an assumption, and alternative ones are also defensible.[26] For instance, central banks can judge that they are more unsure of the interest rate pass-through when interest rates are low because of potential side effects, and more confident of it when interest rates are high. In a world where the secular, steady-state level of natural interest rates is low, this means that the least uncertain policy $\underline{r}$ is above steady state.

Besides, it can be argued that the policy whose effects are best understood varies across time. For instance, it can be at the level of the policy rates that have been *recently* implemented, which may not correspond to the steady state. Such an assumption is implicit in the reliance on the Brainard principle to justify gradualism, or the terminology of *"conservatism principle"* used by Blinder (1999). For instance, in March 2019, at the time of the quote by Mario Draghi mentioned in the introduction, the EONIA (euro overnight index average) had been at levels below 1.5 percent for more than 10 years, far below its long-term average. In this context, moving cautiously may be better interpreted as tilting nominal rates toward recent low levels, not toward their historical average—which would mean a sharp and sudden increase in rates.

In this section we generalize the setup of Section 2 to allow for the possibility that the policy whose effects are best understood $\underline{r}$ is not the steady-state level of natural rates. We allow for two generalizations. First, we allow $\underline{r}$ to be any arbitrary function of the past. For instance, it can be a weighted average of recent past real interest rates, as would be the case if the central bank is more confident

[26]The importance of the reference point of minimal uncertainty was not lost to Brainard, who cautioned that "some care must be used in interpreting [the attenuation principle]. The gap in this context is not the difference between what policy was 'last period' and what would be required to make the expected value of [the target variable] equal to [its target]. In the example we have used, the gap is the difference between the point where the variance of [the target variable] is least and [the policy instrument] required to give an expected value of [the target variable] equal to [its target]" (Brainard 1967). Sack (1998) and Wieland (2000) embed the Brainard principle in a dynamic learning model to capture the interpretation of the Brainard principle as a recommendation for gradualism.

about the effect of the policies it recently implemented. Second, its unconditional average can differ from the steady-state level of interest rates. We add the time index -1 to $\underline{r}_{-1}$ to emphasize that it is measurable with information available at $t-1$. Equation (7) is now

$$\pi = -\phi(r - \underline{r}_{-1}) - \bar{\phi}\underline{r}_{-1} + \varepsilon + E_{-1}(\pi). \tag{41}$$

The constant term $-\bar{\phi}\underline{r}_{-1}$ is necessary to guarantee that the central bank's average expectation of inflation across all the models it considers is correct, $E^*(\pi) = -\bar{\phi}r + \varepsilon + E_{-1}(\pi)$.

The program of the central bank is still to minimize the loss function (6), which can still be decomposed into a mean term (9) and a variance term. The only difference relative to Section 2 is that, by definition, the variance of inflation

$$Var^*(\pi) = V_\phi(r - \underline{r}_{-1})^2 \tag{42}$$

is now minimized for $r = \underline{r}_{-1}$. The central bank still solves the trade-off between its two objectives by setting the interest rate r to a midpoint between the interest rate (11) that minimizes the mean term (9) and the interest rate $\underline{r}$ that minimizes the variance term (42):

$$r = \alpha r^s + (1-\alpha)\underline{r}_{-1}, \tag{43}$$

where r^s is still given by (11) and α is still given by (14).

5.2 Missing the Inflation Target on Average

Because the real interest rate does not track the natural rate one-for-one, inflation is not on target, which is anticipated by the private sector. Its rational expectations of inflation are

$$E_{-1}(\pi) = \pi^* + \left(\frac{1}{\alpha} - 1\right)\bar{\phi}(E_{-1}(\bar{r}^n) - \underline{r}_{-1}). \tag{44}$$

It follows in particular that unconditional average inflation is

$$E(\pi) = \pi^* - \left(\frac{1}{\alpha} - 1\right)\bar{\phi}E(\underline{r}_{-1}) \neq \pi^*. \tag{45}$$

Only when $E(\underline{r}_{-1}) = 0$ is average inflation on target π^*. If the central bank understands better how its policy affects inflation around a rate $\underline{r}_{-1} > 0$ which is on average greater that the steady-state natural rate, Brainard's attenuation principle provides an argument for setting the real interest rate above the natural interest rate on average. As a consequence, average inflation is below target π^* on average. Conversely, if the central bank understands better how the economy works around a rate $\underline{r}_{-1} < 0$ lower that the steady-state natural rate, average inflation is above target π^* on average.

Having average inflation not equal to π^* could be desirable, since it could come with the benefit of less uncertain inflation. In order to assess whether the desire to let average inflation depart from π^* constitutes a bias, Appendix B solves for the optimal average inflation rate under commitment to show that it does constitute a bias.[27]

PROPOSITION 6 (A Cautiousness Bias on Average Inflation). *Regardless of the value of* $\underline{r}_{-1}$*, the optimal average inflation rate is* π^**. When* $E(\underline{r}_{-1}) \neq 0$*, concerns about uncertainty make average inflation depart from this optimal inflation target.*

Therefore, the departure of average inflation from π^* when $E(\underline{r}_{-1}) \neq 0$ is indeed a second manifestation of the cautiousness bias, this time on average inflation. Because the desire of the central bank to systematically tilt the real interest rate toward $E(\underline{r}_{-1})$ is fully anticipated by the private sector, in equilibrium the central bank fails to do so and the real interest rate is still (17). The desire to tilt the interest rate toward $E(\underline{r}_{-1})$ only results in an inflationary bias (if $E(\underline{r}_{-1}) < 0$) or deflationary bias (if $E(\underline{r}_{-1}) > 0$).

As a result, in the generic case when the best-understood policy is not on average the steady-state policy $E(\underline{r}_{-1}) \neq 0$, a discretionary central bank cannot follow a cautious strategy without failing to deliver on its inflation mandate on average inflation. Only in the particular case when the best-understood policy is on average the

[27]The inflation target is usually understood as the average inflation rate desired by the central bank, i.e., the optimal average inflation rate under commitment. Therefore, by referring to π^* as the central bank's inflation target, we have implicitly already assumed that π^* is the average inflation rate under commitment. Appendix B shows it is indeed the case. It is not simply by definition of π^*, however: what the reduced-form preferences (6) assume by construction is only that π^* minimizes the mean term of the loss function. With Brainard uncertainty, another average inflation rate could minimize the variance term if the central bank were able to affect the average real interest rate. But the Phillips-curve constraint (2) imposes that it cannot.

steady-state policy $E(\underline{r}_{-1}) = 0$ can a cautious discretionary central bank deliver an average inflation rate in line with its inflation mandate. In this case, inflation is still off target more often than if the central bank were not cautious, but symmetrically so.

5.3 No Impact on Equilibrium Interest Rates

While allowing the least uncertain policy $\underline{r}_{-1}$ to differ from the steady-state level of natural rates affects the equilibrium level of inflation, it makes no change to the equilibrium level of real interest rates. Indeed, plugging the private sector's expectations of inflation in equilibrium (44) into the expression for the policy rate chosen by the central bank (13), we find that the equilibrium real interest rate is still given by Equation (17).

Because under the New Classical Phillips curve the private sector observes $\underline{r}_{-1}$, any change that $\underline{r}_{-1}$ makes to the central bank's policy is anticipated by the private sector and fails to affect equilibrium real interest rates—it only affects inflation. In particular, time variations in $\underline{r}_{-1}$, as would occur if the central bank gradually adjusts its $\underline{r}_{-1}$ to recent levels of the real interest rate, do not generate any persistence in real interest rates, as Equation (43) could at first suggest.

6. Conclusion

Since Alan Blinder's book (Blinder 1999) made Brainard's attenuation principle widely known to central bankers, the economic literature has found several instances in which the Brainard principle proved not robust, with uncertainty calling instead for a more aggressive policy response. Preempting the literature to come—and because Brainard's original paper already emphasized cases in which uncertainty called for policy aggressiveness—Blinder commented: "My intuition tells me that [Brainard's principle] is more general—or at least more wise—in the real world than the mathematics will support."

In this paper, we made a distinct qualification to Brainard's attenuation principle. Focusing on situations in which uncertainty does rationalize policy attenuation, we showed that, when policy outcomes depend on the expectations of the private sector as in monetary policy, the desire to attenuate policy can backfire. It adversely shifts the private sector's inflation expectations, forcing the central bank to ultimately act by as much, but for worse outcomes. Our analysis does not conclude that uncertainty does not justify moving cautiously. But it emphasizes that central banks face a bias toward being *overly* cautious.

Appendix A. Microfoundations of the Shocks to the Euler Equation

In this appendix, we discuss the connection between the alternative representations of the shocks to the Euler equation: through shocks to the underlying fundamentals such as productivity a_t, through shocks to natural output y_t^n, through shocks to the variable $v_t = y_t^n - E_t(y_{t+1}^n) = -\sigma r_t^n$, or through shocks to the natural rate r_t^n. We show that when the central bank faces uncertainty about σ, only the first three are equivalent.

To do so, we first rederive the Euler equation (1) through a standard microfounded model with no capital and technology shocks as the only fundamental disturbance. A representative household consumes C_t, works L_t hours, and invests in B_t nominal riskless bonds in order to maximize intertemporal utility:

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \frac{L_t^{1+\psi}}{1+\psi} \right), \tag{A.1}$$

where σ is the intertemporal elasticity of substitution, ψ is the inverse of the Frisch elasticity of labor, and $\beta \in (0, 1)$ is the household's discount factor.

A unit of consumption costs the price P_t. A unit of labor is paid the nominal wage W_t. The household chooses to invest B_t in nominal riskless bonds yielding the nominal interest rate I_t. The household receives nominal profits Ω_t from firms. It faces the flow budget constraint:

$$P_t C_t + B_t = W_t L_t + \Omega_t + I_t B_{t-1} \tag{A.2}$$

and an additional borrowing constraint that prevents it from entering Ponzi schemes. The household takes all prices as given.

Its optimal labor supply decision is to equate its marginal rate of substitution to the real wage $w_t = W_t/P_t$:

$$L_t^{\psi} C_t^{\frac{1}{\sigma}} = w_t. \tag{A.3}$$

The household's total consumption C_t results from the consumption C_t^i of a continuum $i \in [0, 1]$ of individual goods. We assume standard CES preferences with an elasticity of substitution across goods θ. The household's demand for good i is therefore

$$C_t^i = \left(\frac{P_t^i}{P_t}\right)^{-\theta} C_t. \tag{A.4}$$

Firm i produces good i using the production function:

$$Y_t^i = A_t (N_t^i)^{\alpha}, \tag{A.5}$$

where A_t is an aggregate productivity shock. Under flexible prices, firm i sets its price P_t^i to maximize present-period profits, internalizing the demand curve (A.4) it faces. It charges a markup over marginal costs:

$$P_t^i = \frac{\theta}{\theta - 1} P_t \frac{w_t}{A_t \alpha (N_t^i)^{\alpha - 1}}. \tag{A.6}$$

Define natural output as the value of output in a flexible-price equilibrium. In a flexible-price equilibrium all firms set the same price and $A_t \alpha N_t^{\alpha-1} = \frac{\theta}{\theta-1} w_t$, where N_t is total labor demanded by firms. Combining the first-order conditions of the household and the firms and assuming that the goods and labor markets clear, $C_t = Y_t$ and $L_t = N_t$, gives natural output as a function of technology. Using lowercase variables to denote log-deviations from a steady state with $A = 1$, it is given by

$$y_t^n = \frac{\psi + 1}{1 + \psi + \left(\frac{1}{\sigma} - 1\right) \alpha} a_t. \tag{A.7}$$

Natural output is a function of the exogenous shock a_t. Note that natural output depends on the parameter σ, but only because the standard preferences we have assumed make σ parameterize both the intertemporal elasticity of substitution and the income effect on labor supply. Natural output depends on the strength of the income effect but not on the intertemporal elasticity of substitution. What we assume to be uncertain for the central bank is the intertemporal elasticity of substitution, not the strength of the income effect. Therefore, we assume that model uncertainty does not affect the central bank's expectations of natural output.[28]

The household's investment decision results in the Euler equation. Taking into account goods market clearing $C_t = Y_t$, it becomes in log-linear form:

$$y_t = -\sigma(i_t - E_t(\pi_{t+1})) + E_t(y_{t+1}). \tag{A.8}$$

The Euler equation applies in particular under flexible prices, in which case it residually gives the real interest rate in the flexible-price equilibrium, that is, the natural rate:

$$r_t^n = -\frac{1}{\sigma}(y_t^n - E_t(y_{t+1}^n)). \tag{A.9}$$

To rewrite the Euler equation in terms of difference with respect to natural output, define the output gap $x_t \equiv y_t - y_t^n$. The Euler equation is

$$x_t = -\sigma(i_t - E_t(\pi_{t+1})) + E_t(x_{t+1}) + v_t, \tag{A.10}$$

$$v_t \equiv -(y_t^n - E_t(y_{t+1}^n)). \tag{A.11}$$

The disturbance v_t is a function of natural output, therefore of the exogenous shocks. Since $r_t^n = \frac{1}{\sigma}v_t$, in models where agents face no model uncertainty it is customary to express the shock v_t in the Euler equation as exogenous variations in the natural rate:

$$x_t = -\sigma(i_t - E_t(\pi_{t+1}) - r_t^n) + E_t(x_{t+1}). \tag{A.12}$$

[28]There are several ways to make the parameter σ play the role of the elasticity of substitution only, and therefore to explicitly eliminate the dependence of natural output on σ. For instance, we could assume GHH preferences (Greenwood, Hercowitz, and Huffman 1988) to eliminate the income effect on labor supply. Alternatively, we could disentangle the intertemporal elasticity of substitution and the income effect on labor supply through Epstein-Zin preferences (Epstein and Zin 1989). However, in both cases the Euler equation would slightly differ from its canonical form. We thus stick to the standard preferences.

However, the two representations (A.10) and (A.12) are not equivalent when the central bank faces uncertainty over σ. The variables a_t, y_t^n, and v_t are independent of the intertemporal elasticity of substitution, while σ enters the definition (A.9) of r_t^n. Parameterizing the shocks to the Euler equation through an exogenous distribution for r_t^n in Equation (A.12) would spuriously make the effect of disturbances appear dependent on the value of σ, whereas σ multiplies r_t^n in Equation (A.12) only because r_t^n is divided by σ in definition (A.9). While the issue is irrelevant in models without parameter uncertainty, it matters when the central bank faces uncertainty on σ, because it changes the value of i_t for which the variance of x_t is minimal.

Appendix B. Proofs of Propositions 2 and 6: Optimal Policy Under Commitment

When the central bank sets policy under commitment, it understands the effect of its policy on the inflation expectations of the private sector. Because it understands that the private sector forms rational expectations in accordance with (7), it understands that in equilibrium its policy r must satisfy

$$E_{-1}(r) = E_{-1}(r^n). \tag{B.1}$$

Because the constraint (B.1) on policy rates spreads across realizations of ε, the program of the central bank no longer reduces to independent programs for each realization of ε. Instead, the central bank's faces one program at each information node of the private sector. At each node, it chooses the policy rates $r(\varepsilon)$ in each final realization of the shock, and the unique expectation of the private sector e to minimize:

$$\min_{(r(\varepsilon))_\varepsilon, e} E_{-1}(\mathcal{L}(\varepsilon))$$

$$= E_{-1}\left(\left(-\bar{\phi}(r(\varepsilon) - r^n(\varepsilon)) + e - \pi^*\right)^2 + V_\phi r(\varepsilon)^2\right), \tag{B.2}$$

$$\text{s.t. } E_{-1}(r(\varepsilon)) = E_{-1}(r^n(\varepsilon)). \tag{B.3}$$

Noting 2λ the Lagrange multiplier on the constraint, the first-order conditions (FOCs) are

$$\forall\varepsilon, /r(\varepsilon) : \bar{\phi}^2\left(r(\varepsilon) - r^n(\varepsilon) - \frac{e-\pi^*}{\bar{\phi}}\right) + V_\phi r(\varepsilon) + \lambda = 0, \quad \text{(B.4)}$$

$$/e : e = E_{-1}\left(\bar{\phi}(r(\varepsilon) - r^n(\varepsilon)) + \pi^*\right). \quad \text{(B.5)}$$

Using the constraint (B.3), the FOC (B.5) gives $e = \pi^*$. Inflation expectations are always on target. Taking expectations E_{-1} of the FOC (B.4) and using the constraint (B.3) solves for λ. Substituting the expression for λ in the FOC solves for $r(\varepsilon)$:

$$r(\varepsilon) = E_{-1}(r^n(\varepsilon)) + \alpha(r^n(\varepsilon) - E_{-1}(r^n(\varepsilon))). \quad \text{(B.6)}$$

The policy rate takes the exact same value as under discretion (17). Substituting the policy rate (B.6) into Equation (7) gives the departure of inflation from target under commitment (19) in Proposition 2.

In Section 5, we generalize the setup of Section 2 by replacing Equation (7) with Equation (41). The optimal policy under commitment keeps setting expected inflation on target $E_{-1}(\pi) = \pi^*$ in this case, and therefore unconditional average inflation on target, $E(\pi) = \pi^*$. Indeed, the only difference with respect to the case $\underline{r}_{-1} = 0$ is to replace the first-order condition (B.4) with

$$\forall\varepsilon, /r(\varepsilon) : \bar{\phi}^2\left(r(\varepsilon) - r^n(\varepsilon) - \frac{e-\pi^*}{\bar{\phi}}\right) + V_\phi(r(\varepsilon) - \underline{r}_{-1}) + \lambda = 0. \quad \text{(B.7)}$$

Equation (B.5) is unchanged. Using the constraint (B.3), it still gives $e = \pi^*$. Following the same steps as in the case $\underline{r} = 0$, one can also check that the policy rate still takes the value (B.6) in this generalized case, as it does under discretion. This proves Proposition 6.

Appendix C. Proof of Proposition 3: Optimal Discounting of Uncertainty Concerns

A central banker that puts a weight $\delta \geq 0$ on the variance term in the loss function (20) sets the interest rate to $r^d = \alpha r^s$, where

$\alpha = \frac{\bar{\phi}^2}{\bar{\phi}^2+\delta V_\phi}$. As δ increases from zero to infinity, α decreases from 1 to 0: the more concerned he is about uncertainty, the less he reacts to shocks. We can therefore parameterize the central banker's type by how aggressively he reacts to shocks, as captured by α. An α-type central banker acting under discretion delivers an inflation rate of

$$E^*(\pi) - \pi^* = (1-\alpha)\bar{\phi}\left(\frac{1}{\alpha}E_{-1}(r^n) + (r^n - E_{-1}(r^n))\right). \quad \text{(C.1)}$$

Society compares these outcomes using its loss function with $\delta = 1$. The average loss generated by an α-type central banker is

$$E[\mathcal{L}(\varepsilon)] = Var(E^*(\pi) - \pi^*) + Var^*(\pi). \quad \text{(C.2)}$$

The two terms can be written as

$$Var(E^*(\pi) - \pi^*) = (1-\alpha)^2\bar{\phi}^2\left(\frac{1}{\alpha^2}V_E + V_U\right), \quad \text{(C.3)}$$

$$Var^*(\pi) = V_\phi(V_E + \alpha^2 V_U), \quad \text{(C.4)}$$

where $V_E \equiv Var(E_{-1}(r^n))$ is the variance of fluctuations in the natural rate that are expected by the private sector, and $V_U = Var(r^n - E_{-1}(r^n))$ is the variance of fluctuations in the natural rate that are unexpected by the private sector. Therefore, society wants to appoint the central banker whose α minimizes

$$E[\mathcal{L}(\varepsilon)] = \left((1-\alpha)^2\bar{\phi}^2 + \alpha^2 V_\phi\right)\left(\frac{1}{\alpha^2}V_E + V_U\right). \quad \text{(C.5)}$$

Taking the log and differentiating in α, the optimal α satisfies

$$\frac{\alpha V_\phi - (1-\alpha)\bar{\phi}^2}{(1-\alpha)^2\bar{\phi}^2 + \alpha^2 V_\phi} = \frac{1}{\alpha + \alpha^3\frac{V_U}{V_E}}. \quad \text{(C.6)}$$

The right-hand-side term is decreasing from infinity to $V_E/(V_E+V_U)$ as α increases from 0 to 1. Define the left-hand-side term as the function f:

$$f(\alpha) = \frac{\alpha - \alpha^*}{(1-\alpha)^2\alpha^* + \alpha^2(1-\alpha^*)}, \quad \text{(C.7)}$$

where $\alpha^* = \frac{\bar{\phi}^2}{\bar{\phi}^2+V_\phi}$ is the value of α of the central banker who has the same preferences as society, $\delta = 1$. The LHS f is negative for $\alpha < \alpha^*$, so it can only cross the RHS over $[\alpha^*, 1]$. The derivative of f has the sign of the quadratic polynomial $P(\alpha) = -\alpha^2 + 2\alpha^*\alpha + \alpha^*(1 - 2\alpha^*)$. The polynomial reaches its maximum at $\alpha = \alpha^*$ and has two real roots. If $\alpha^* > 1/2$, the larger root is greater than 1, so P is positive over $[\alpha^*, 1]$. It follows that f is increasing over $[\alpha^*, 1]$. There is a unique crossing of the RHS and LHS terms in Equation (C.6). If $\alpha^* < 1/2$, then the second root is smaller than 1, so f is increasing then decreasing over $[\alpha^*, 1]$. Yet, since $f(1) = 1 > V_E/(V_E + V_U)$, there is still a unique crossing of the RHS and LHS terms in Equation (C.6). In both cases, the two curves cross at a value greater than α^*, unless $V_E = 0$, in which case the RHS is constantly equal to zero and crosses the LHS at zero. An increase in $V_E/(V_E + V_U)$ causes the RHS to shift up: the optimal α therefore increases with the fraction of shocks that are expected by the private sector.

Appendix D. Proof in the Case of a Dual Mandate

We generalize the case of a single inflation mandate considered in the main text of Section 2 to allow for a dual objective to stabilize both inflation and the output gap. In doing so, we also allow for the possibility that the central bank seeks to set output above potential $x^* > 0$, which will result in an inflation bias. The present-period loss of the central bank is

$$\mathcal{L}(\varepsilon) = E^*((\pi - \pi^*)^2) + \lambda E^*((x - x^*)^2), \tag{D.1}$$

where λ is the preference weight of the central bank on stabilizing the output gap. The appearance of the real interest rate in the determination of the output gap,

$$x = -\sigma r + v, \tag{D.2}$$

and inflation,

$$\pi = \kappa(-\sigma r + v) + e(\pi), \tag{D.3}$$

are unchanged. The mean squared errors in the loss function (D.1) can still be decomposed into a term for squared distances to targets and a term for variances:

$$\mathcal{L}(\varepsilon) = \Big((E^*(\pi) - \pi^*)^2 + \lambda(E^*(x) - x^*)^2 \Big) + \Big(Var^*(\pi) + \lambda Var^*(x) \Big), \tag{D.4}$$

$$\text{where } (E^*(\pi) - \pi^*)^2 + \lambda(E^*(x) - x^*)^2 = (-\kappa\bar{\sigma}r + \varepsilon + e(\pi) - \pi^*)^2 + \lambda(-\bar{\sigma}r + v - x^*)^2, \tag{D.5}$$

$$\text{and } Var^*(\pi) + \lambda Var^*(x) = (\kappa^2 + \lambda)V_\sigma r^2. \tag{D.6}$$

Denote r^s the interest rate that the central bank sets when it faces no model uncertainty, $V_\sigma = 0$. In this case the central bank can focus on minimizing the first term (D.5) in its loss function. It sets

$$r^s = r^n + \frac{\kappa}{\bar{\sigma}(\kappa^2 + \lambda)}(e(\pi) - \pi^*) - \frac{\lambda}{\bar{\sigma}(\kappa^2 + \lambda)}x^*. \tag{D.7}$$

A desire to stabilize the output gap $\lambda > 0$ changes the optimal discretionary policy relative to the case in which the central bank has no concerns about uncertainty in two ways. First, the central bank reacts less to departures of inflation expectations from target. This is regardless of whether the central bank seeks to set output above potential $x^* > 0$. Second, when $x^* > 0$ the central bank seeks to set the interest rate lower in order to set the output gap higher. This last feature of the discretionary policy results in Kydland and Prescott (1977) and Barro and Gordon (1983a, 1983b)'s inflation bias. Rational expectations of inflation are above target:

$$E_{-1}(\pi) = \pi^* + \frac{\lambda}{\kappa}x^* > \pi^*, \tag{D.8}$$

but the output gap is $x = 0$.[29]

When the central bank is uncertain about the impact of its rate decision on inflation and the output gap, $V_\sigma > 0$, the policy rate

[29]Monetary policy could surprise the private sector by responding to unexpected shocks to the natural interest rate, which would make the output gap depart from zero (although it would need to be zero on average). The central bank has no interest in doing so here, however, because there are no cost-push shocks.

r also affects the variance term (10) in the loss function. The central bank solves this problem by choosing a midpoint r between the optimal interest rate policy without model uncertainty r_s which minimizes the first term, and the steady-state interest rate $r = 0$ which minimizes the second term:

$$r = \alpha r^s, \tag{D.9}$$

where α is still given by (14).

We solve for the rational expectations that this policy generates. Injecting policy (D.9) into Equation (7), taking expectations E_{-1}, and imposing rational expectations $e(\pi) = E_{-1}(\pi)$ yields

$$E_{-1}(\pi) = \pi^* + \left(\frac{1}{\alpha} - 1\right) \frac{\bar{\sigma}(\kappa^2 + \lambda)}{\kappa} E_{-1}(r^n) + \frac{\lambda}{\kappa} x^*. \tag{D.10}$$

Inflation expectations can deviate from target for two reasons. First is the inflation bias, as noted in the case in which there are no concerns about uncertainty: a desire to set output above potential $\lambda > 0$ results in higher inflation expectations. Second is the cautiousness bias: a concern about parameter uncertainty $\alpha < 1$ results in lower (higher) inflation expectations when the natural rate is below (above) steady state. The generalization to the case of a dual mandate shows both that the cautiousness bias is robust to a dual mandate, and that the cautiousness bias and inflation bias arise from distinctly different perverse incentives.

Plugging expectations (D.10) into the optimal policy rate (D.9), the expression for the real interest rate is exactly the same as (17) under a single inflation mandate. In equilibrium the central bank attenuates its response only to unforeseen changes in the natural rate. It does so in exactly the same proportions as in the case of a single inflation mandate.

Appendix E. Proof of Proposition 4: Dynamics of the System and IRF in the Case of the Sticky-Information Phillips Curve

Injecting the solution (33) for the long-term interest rate into the Euler equation (24) (of the true model) gives the output gap as

$$x_t = \bar{\sigma}(1 - \alpha) R_t^n - \frac{\alpha}{\kappa} \bar{E}_{t-1}(\pi_t - \pi^* + \zeta \Delta x_t). \tag{E.1}$$

Using the sticky-information Phillips curve (22) to replace the last expectation term gives Equation (34). Equation (35) is simply the sticky-information Phillips curve.

We solve for the IRF for a natural rate shock using the method of undetermined coefficients, following Mankiw and Reis (2007). Under the assumption of an AR(1) process for the natural rate r_t^n, and denoting $\hat{\pi}_t = \pi_t - \pi^*$ the deviation of inflation from its target, the dynamic system is

$$x_t = \frac{\bar{\sigma}}{1-\rho} r_t^n - \frac{\alpha}{(1-\alpha)\kappa}\hat{\pi}_t, \tag{E.2}$$

$$\hat{\pi}_t = \kappa x_t + \bar{E}_{t-1}(\hat{\pi}_t + \zeta \Delta x_t). \tag{E.3}$$

Under the assumption of an AR(1) process for the natural rate r_t^n, its Wold decomposition is

$$r_t^n = \sum_{k=0}^{\infty} \rho^k \eta_{t-k}. \tag{E.4}$$

We look for a solution where $\hat{\pi}_t$ and x_t are functions of the fundamental shock only. We write their Wold decompositions:

$$\hat{\pi}_t = \sum_{k=0}^{\infty} \phi_k^{\pi} \eta_{t-k}, \tag{E.5}$$

$$x_t = \sum_{k=0}^{\infty} \phi_k^{x} \eta_{t-k}, \tag{E.6}$$

with coefficients $(\phi_k^{\pi})_k$ and $(\phi_k^{x})_k$ to be determined. To translate the dynamic system (E.2)–(E.3) into equations in ϕ_k^{π} and ϕ_k^{x}, note that

$$\begin{aligned}
&\bar{E}_{t-1}(\pi_t + \zeta \Delta x_t) \\
&\quad = \sum_{k=1}^{\infty} \left(1-(1-\lambda)^k\right)\left(\phi_k^{\pi} + \zeta(\phi_k^{x} - \phi_{k-1}^{x})\right)\eta_{t-k}.
\end{aligned} \tag{E.7}$$

The dynamic system (E.2)–(E.3) therefore can be written as

$$\sum_{k=0}^{\infty} \phi_k^x \eta_{t-k} = \frac{\bar{\sigma}}{1-\rho} \sum_{k=0}^{\infty} \rho^k \eta_{t-k} - \frac{\alpha}{(1-\alpha)\kappa} \sum_{k=0}^{\infty} \phi_k^\pi \eta_{t-k}, \tag{E.8}$$

$$\sum_{k=0}^{\infty} \phi_k^\pi \eta_{t-k} = \kappa \sum_{k=0}^{\infty} \phi_k^x \eta_{t-k} + \sum_{k=1}^{\infty} \Big(1-(1-\lambda)^k\Big)\Big(\phi_k^\pi + \zeta(\phi_k^x - \phi_{k-1}^x)\Big)\eta_{t-k}. \tag{E.9}$$

Identifying the coefficients, it implies the following difference equations in $(\phi_k^\pi)_k$ and $(\phi_k^x)_k$:

$$\forall k \geq 0, \phi_k^x = \frac{\bar{\sigma}}{1-\rho}\rho^k - \frac{\alpha}{(1-\alpha)\kappa}\phi_k^\pi, \tag{E.10}$$

$$\forall k \geq 1, \phi_k^\pi = \kappa\phi_k^x + \Big(1-(1-\lambda)^k\Big)\Big(\phi_k^\pi + \zeta(\phi_k^x - \phi_{k-1}^x)\Big), \tag{E.11}$$

$$\text{for } k = 0, \phi_0^\pi = \kappa\phi_0^x. \tag{E.12}$$

Using Equation (E.11) to eliminate ϕ_k^π in (E.10) gives the following first-order difference equation in ϕ_k^x:

$$\forall k \geq 1, \left((1-\lambda)^k + \frac{\alpha}{1-\alpha}\left(1+\frac{\zeta}{\kappa}\Big(1-(1-\lambda)^k\Big)\right)\right)\phi_k^x = \left(\frac{\alpha\zeta}{(1-\alpha)\kappa}\Big(1-(1-\lambda)^k\Big)\right)\phi_{k-1}^x + \frac{\bar{\sigma}}{1-\rho}\rho^k(1-\lambda)^k. \tag{E.13}$$

This gives ϕ_k^x as a function of ϕ_{k-1}^x. We obtain the entire sequence of $(\phi_k^x)_k$ from the initial condition $\phi_0^x = \frac{(1-\alpha)\bar{\sigma}}{1-\rho}$. We then recover the solution for inflation from (E.11). The solutions for the interest rates follow.

Appendix F. Proof of Proposition 5: Derivation in the Case of the NKPC

Since the central bank under discretion takes the future policy rates $E_t(R_{t+1})$ as given, the central bank equivalently chooses the long-term real interest rate $R_t = r_t + E_t(R_{t+1})$. The long-term real interest rate that sets inflation on target is

$$R_t^s = R_t^n + \frac{\beta E_t(\pi_{t+1})}{\bar{\phi}}. \tag{F.1}$$

The long-term real interest rate that minimizes the within variance of inflation is $R_t = 0$. Taking the first-order condition of the loss function $\mathcal{L}_t = E_t^*(\pi_t^2)$ where π_t satisfies (38) gives that the central bank sets its policy rate so that the long-term rate is equal to the weighted average (39) of these two rates. Injecting the long-term rate into (38) (for the true model) gives inflation as

$$\pi_t = (1-\alpha)\left(\bar{\phi}R_t^n + \beta E_t(\pi_{t+1})\right) \tag{F.2}$$

or, written with time-series polynomials,

$$\pi_t = \Big(I - \beta(1-\alpha)F\Big)^{-1}(1-\alpha)\bar{\phi}R_t^n, \tag{F.3}$$

where I is the identity polynomial and F is the forward polynomial.

Under the assumption that r_t^n (and therefore R_t^n) follows an AR(1) process (36), the solution to (F.2) is

$$\pi_t = \frac{(1-\alpha)\bar{\phi}}{1-\beta(1-\alpha)\rho}R_t^n. \tag{F.4}$$

This implies that the private sector forms expectations of inflation:

$$E_t(\pi_{t+1}) = \frac{(1-\alpha)\bar{\phi}}{1-\beta(1-\alpha)\rho}\rho R_t^n. \tag{F.5}$$

Injecting these inflation expectations into the solution for the long-run rate (39) gives (40). The short-term rate $r_t = R_t - E_t(R_{t+1})$ is similarly

$$r_t = \alpha\left(\frac{1}{1-\beta(1-\alpha)\rho}\right)r_t^n. \tag{F.6}$$

It follows that the ex ante nominal long-term interest rate $I_t = R_t + E_t\left(\sum_{k=0}^{\infty}\pi_{t+k+1}\right)$ is

$$I_t = \left(\frac{\alpha + (1-\alpha)\bar{\phi}\frac{\rho}{1-\rho}}{1-\beta(1-\alpha)\rho}\right)R_t^n, \tag{F.7}$$

whereas it is $I_t = R_t^n$ in the absence of concerns about uncertainty. The coefficient in front of R_t^n tends toward infinity as ρ tends toward 1. Therefore, for persistent enough shocks, the central bank ends up moving the nominal long-term interest rate by more than it would have in the absence of concerns about uncertainty. Since the nominal short-term interest rate $i_t = r_t + E_t(\pi_{t+1})$ is similarly

$$i_t = \left(\frac{\alpha + (1-\alpha)\bar{\phi}\frac{\rho}{1-\rho}}{1-\beta(1-\alpha)\rho} \right) r_t^n, \tag{F.8}$$

the same conclusion applies to the nominal short-term rate.

Appendix G. The Case of the NKPC and the Recursive Euler Equation

Plugging the recursive Euler equation (24) into the NKPC (37), the relationship between the short-term interest rate chosen at t and inflation at t is

$$\pi_t = -\phi r_t + \kappa E_t(x_{t+1}) + \kappa v_t + \beta E_t(\pi_{t+1}), \tag{G.1}$$

where $\phi = \kappa\sigma$. Because the central bank acts under discretion, it chooses r_t at t taking $E_t(x_{t+1})$ and $E_t(\pi_{t+1})$ as given. It does so to minimize the loss $\mathcal{L}_t = E_t^*(\pi_t^2)$. The interest rate that sets inflation on target is

$$r_t^s = r_t^n + \frac{E_t(x_{t+1})}{\bar{\sigma}} + \frac{\beta E_t(\pi_{t+1})}{\bar{\phi}}. \tag{G.2}$$

The interest rate that minimizes the within variance of inflation is $r_t = 0$. The central bank sets its policy rate to the weighted average of these two rates:

$$r_t = \alpha r_t^s. \tag{G.3}$$

Injecting this solution for the short-term interest rate into (G.1) and using the NKPC (37) to replace future output gap $E_t(x_{t+1})$ gives inflation (in the true model) as

$$\pi_t = (1-\alpha)\Big(\bar{\phi} r_t^n + (1+\beta)E_t(\pi_{t+1}) - \beta E_t(\pi_{t+2}) \Big). \tag{G.4}$$

This is a second-order difference equation in π_t. Noting F the forward time-series operator, the stationary solution is

$$\pi_t = \left[I - (1-\alpha)(1+\beta)F + (1-\alpha)\beta F^2 \right]^{-1} (1-\alpha)\bar{\phi} r_t^n. \quad \text{(G.5)}$$

Under the assumption that r_t^n follows an AR(1) process (36), the solution to (G.5) is

$$\pi_t = \frac{(1-\alpha)\bar{\phi}}{1-(1-\alpha)(1+\beta)\rho + (1-\alpha)\beta\rho^2} r_t^n. \quad \text{(G.6)}$$

Expectations of inflation $E_t(\pi_{t+1})$ and of the output gap $E_t(x_{t+1})$ follow. Plugging them into the expression for the short-term interest-rate (G.3) gives the solution for the short-term interest rate:

$$r_t = \alpha \left(\frac{1}{1-(1-\alpha)\rho(1+\beta(1-\rho))} \right) r_t^n. \quad \text{(G.7)}$$

As in the case of the iterated Euler equation, Brainard's attenuation principle leads the central bank to move rates by less by a factor $\alpha < 1$, but the reaction of inflation expectations forces the central bank to move rates by more, this time by a factor $1/(1-(1-\alpha)\rho(1+\beta(1-\rho))) > 1$. Because the more persistent a shock is, the more inflation expectations react, persistent shocks force the central bank to react the most vigorously. The central bank varies the real interest rate by *more* than the natural rate if and only if the coefficient on r_t^n in (G.7) is greater than 1. This happens if and only if the second-order polynomial,

$$P(\rho) = \beta\rho^2 - (1+\beta)\rho + 1, \quad \text{(G.8)}$$

takes negative values. Because the roots of the polynomial are 1 and $1/\beta$, this never happens for $\rho \in [0,1]$. However, as in the case of the iterated Euler equation, the short-term nominal interest rate can vary more than if the central bank has no concerns over Brainard uncertainty. Indeed, the solution for the short-term nominal interest rate $i_t = r_t + E_t(\pi_{t+1})$ is

$$i_t = \left(\frac{\alpha + (1-\alpha)\bar{\phi}\rho}{1-(1-\alpha)\rho(1+\beta(1-\rho))} \right) r_t^n. \quad \text{(G.9)}$$

The coefficient in front of r_t^n tends toward $1 + \left(\frac{1-\alpha}{\alpha}\right)\bar{\phi} > 1$ as ρ tends toward 1. Therefore, for persistent enough shocks, the central bank ends up moving the nominal short-term interest rate by more than it would have in the absence of concerns about uncertainty.

References

Acharya, S., J. Bengui, K. Dogra, and S. L. Wee. 2019. "Slow Recoveries and Unemployment Traps: Monetary Policy in a Time of Hysteresis." Working Paper.

Adam, K., and M. Woodford. 2012. "Robustly Optimal Monetary Policy in a Microfounded New Keynesian Model." *Journal of Monetary Economics* 59 (5): 468–87.

Angeletos, G.-M., and J. La'O. 2010. "Noisy Business Cycles." *NBER Macroeconomics Annual 2009*, Vol. 24, ed. D. Acemoglu, K. Rogoff, and M. Woodford, 319–78. NBER, University of Chicago Press.

Ball, L. 1994. "Credible Disinflation with Staggered Price-Setting." *American Economic Review* 84 (1): 282–89.

Barlevy, G. 2011. "Robustness and Macroeconomic Policy." *Annual Review of Economics* 3: 1–24.

Barro, R., and D. Gordon. 1983a. "A Positive Theory of Monetary Policy in a Natural Rate Model." *Journal of Political Economy* 91 (4): 589–610.

———. 1983b. "Rules, Discretion and Reputation in a Model of Monetary Policy." *Journal of Monetary Economics* 12 (1): 101–21.

Bernanke, B. 2007. "Monetary Policy under Uncertainty." Speech given at the 32nd Annual Economic Policy Conference, Federal Reserve Bank of St. Louis, October 19.

Blinder, A. 1999. *Central Banking in Theory and in Practice*. MIT Press.

Brainard, W. 1967. "Uncertainty and the Effectiveness of Policy." *American Economic Review* 57 (2): 411–25.

Carney, M. 2017. "Lambda." Speech given at the London School of Economics, January 16.

Chow, G. 1973. "Effect of Uncertainty on Optimal Control Policies." *International Economic Review* 14 (3): 632–45.

Clarida, R., J. Gali, and M. Gertler. 1999. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature* 37 (4): 1661–1707.

Craine, R. 1979. "Optimal Monetary Policy with Uncertainty." *Journal of Economic Dynamics and Control* 1 (1): 59–83.

Cukierman, A. 1991. "Why Does the Fed Smooth Interest Rates?" In *Monetary Policy on the 75th Anniversary of the Federal Reserve System*, ed. M. Belongia, 111–47. Kluwer Academic Publishers.

Draghi, M. 2019. "Transcripts of the Press Conference of March 7, 2019." European Central Bank.

Eggertsson, G., and M. Woodford. 2003. "Zero Bound on Interest Rates and Optimal Monetary Policy." *Brookings Papers on Economic Activity* 1 (Spring): 139–211.

Epstein, L., and S. Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica* 57 (4): 937–69.

Estrella, A., and F. Mishkin. 1999. "Rethinking the Role of NAIRU in Monetary Policy: Implications of Model Formulation and Uncertainty." In *Monetary Policy Rules,* ed. J. Taylor, 405–36. NBER, University of Chicago Press.

Favaretto, F., and D. Masciandaro. 2016. "Doves, Hawks and Pigeons: Behavioral Monetary Policy and Interest Rate Inertia." *Journal of Financial Stability* 27 (December): 50–58.

Ferrero, G., M. Pietrunti, and A. Tiseno. 2019. "Benefits of Gradualism or Costs of Inaction? Monetary Policy in Times of Uncertainty." Working Paper, Bank of Italy.

Giannoni, M. 2002. "Does Model Uncertainty Justify Caution? Robust Optimal Monetary Policy in a Forward-Looking Model." *Macroeconomic Dynamics* 6 (1): 111–44.

Greenwood, J., Z. Hercowitz, and G. Huffman. 1988. "Investment, Capacity Utilization and the Real Business Cycle." *American Economic Review* 78 (3): 402–17.

Kimura, T., and T. Kirozumi. 2007. "Optimal Monetary Policy in a Micro-founded Model with Parameter Uncertainty." *Journal of Economic Dynamics and Control* 31 (2): 399–431.

Kydland, F., and E. Prescott. 1977. "Rules Rather than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy* 85 (3): 473–92.

Lucas, R. 1972. "Expectations and the Neutrality of Money." *Journal of Economic Theory* 4 (2): 103–24.

Mankiw, G., and R. Reis. 2002. "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* 117 (4): 1295–1328.

———. 2007. "Sticky Information in General Equilibrium." *Journal of the European Economic Association* 5 (2–3): 603–13.

———. 2010. "Imperfect Information and Aggregate Supply." In *Handbook of Monetary Economics*, Vol. 3A, ed. B. Friedman and M. Woodford. Elsevier.

Onatski, A., and J. Stock. 2002. "Robust Monetary Policy under Model Uncertainty in a Small Model of the U.S. Economy." *Macroeconomic Dynamics* 6 (1): 85–110.

Onatski, A., and N. Williams. 2003. "Modeling Model Uncertainty." *European Economic Association* 1 (5): 1087–1122.

Powell, J. 2018. "Monetary Policy in a Changing Economy." Speech given at the Changing Market Structure and Implications for Monetary Policy Economic Symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 24.

Praet, P. 2018. "Assessment of Quantitative Easing and Challenges of Policy Normalisation." Speech given at the ECB and Its Watchers XIX Conference, Frankfurt am Main, March 14.

Reinhart, V. 2003. "Making Monetary Policy in an Uncertain World." In *Monetary Policy and Uncertainty: Adapting to a Changing Economy,* Proceedings of an Economic Policy Symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 28–30.

Riboni, A., and F. Ruge-Murcia. 2010. "Monetary Policy by Committee: Consensus, Chairman Dominance, or Simple Majority." *Quarterly Journal of Economics* 125 (1): 363–416.

Rogoff, K. 1985. "The Optimal Degree of Commitment to an Intermediate Monetary Target." *Quarterly Journal of Economics* 100 (4): 1169–89.

Rudebusch, G. 2001. "Is the Fed Too Timid? Monetary Policy in an Uncertain World." *Review of Economics and Statistics* 83 (2): 203–17.

Sack, B. 1998. "Uncertainty, Learning, and Gradual Monetary Policy." FEDS Working Paper No. 1998-34, Board of Governors of the Federal Reserve System.

———. 2000. "Does the Fed Act Gradually? A VAR Analysis." *Journal of Monetary Economics* 46 (1): 229–56.

Sack, B., and V. Wieland. 2000. "Interest-Rate Smoothing and Optimal Monetary Policy: A Review of Recent Empirical Evidence." *Journal of Economics and Business* 52 (1–2): 205–28.

Sargent, T. 1999. "Comment on "Policy Rules for Open Economies" by Laurence Ball." In *Monetary Policy Rules*, ed. J. Taylor, 144–54. NBER, University of Chicago Press.

Savage, L. 1954. *The Foundations of Statistics.* New York: Dover Publication Inc.

Söderström, U. 2002. "Monetary Policy with Uncertain Parameters." *Scandinavian Journal of Economics* 104 (1): 125–45.

Söderström, U., and K. Leitemo. 2008. "Robust Monetary Policy in the New-Keynesian Framework." *Macroeconomic Dynamics* 12 (S1): 126–35.

Spiegler, R. 2021. "A Simple Model of Monetary Policy under Phillips-Curve Causal Disagreements." Working Paper.

Stein, J., and A. Sunderam. 2018. "The Fed, the Bond Market, and Gradualism in Monetary Policy." *Journal of Finance* 73 (3): 1015–60.

Stock, J. 1999. "Comment on 'Policy Rules for Inflation Targeting' by Glenn Rudebusch and Lars Svensson." In *Monetary Policy Rules,* ed. J. Taylor, 253–59. NBER, University of Chicago Press.

Svensson, L. 1997. "Optimal Inflation Targets, 'Conservative' Central Banks, and Linear Inflation Contracts." *American Economic Review* 87 (1): 98–114.

———. 1999. "Inflation Targeting: Some Extensions." *Scandinavian Journal of Economics* 101 (3): 337–61.

Tetlow, R., and P. von zur Muehlen. 2001. "Robust Monetary Policy with Misspecified Models: Does Model Uncertainty Always Call for Attenuated Policy?" *Journal of Economic Dynamics and Control* 25 (6–7): 911–49.

Theil, H. 1957. "A Note on Certainty Equivalence in Dynamic Planning." *Econometrica* 25 (2): 346–49.

Tillmann, P. 2009. "Optimal Monetary Policy with an Uncertain Cost Channel." *Journal of Money, Credit and Banking* 41 (5): 885–906.

Villeroy de Galhau, F. 2018. "Monetary Policy in the Present Environment: Reconciling Clarity and Flexibility, as Well as Independence and Cooperation." Speech given at the Group of Thirty's annual International Banking Seminar, Bali, Indonesia, October 14.

Walsh, C. 2003. "Implications of a Changing Economic Structure for the Strategy of Monetary Policy." In *Monetary Policy and Uncertainty: Adapting to a Changing Economy,* Proceedings of an Economic Policy Symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 28–30.

Wieland, V. 2000. "Monetary Policy, Parameter Uncertainty and Optimal Learning." *Journal of Monetary Economics* 46 (1): 199–228.

Williams, J. 2013. "A Defense of Moderation in Monetary Policy." *Journal of Macroeconomics* 38 (Part B): 137–50.

Woodford, M. 2003a. "Imperfect Common Knowledge and the Effects of Monetary Policy." In *Knowledge, Information and Expectations in Modern Macroeconomics*, ed. J. S. P. Aghion, R. Frydman, and M. Woodford. Princeton University Press.

———. 2003b. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press.

———. 2003c. "Optimal Interest-Rate Smoothing." *Review of Economic Studies* 70 (4): 861–86.

———. 2010. "Robustly Optimal Monetary Policy with Near-Rational Expectations." *American Economic Review* 100 (1): 274–303.

Monetary Policy Implementation and Payment System Modernization*

Jonathan Witmer
Bank of Canada

24/7 payment settlement may affect the demand for central bank reserves and thus could have an effect on monetary policy implementation. By modifying the standard workhorse model of monetary policy implementation (Poole 1968), we show that 24/7 payment settlement induces a precautionary demand for central bank reserves. Absent any changes or response by the central bank, this will put upward pressure on the overnight interest rate in frameworks with a low level of reserves.

JEL Codes: E, E4, E40, E42, E43.

1. Introduction

Payment, clearing, and settlement systems have undergone drastic changes since banks began accepting claims on each other (Norman, Shaw, and Speight 2011). Technological change and a regulatory interest in systemic risk oversight over the last decade or so has accelerated the pace of these changes. Now, several countries have retail payment systems that provide settlement in real time or near real time 24 hours a day, seven days a week (Tompkins and Olivares 2016). Other countries also plan on adopting such systems. Canada has planned for such a system, and the United States' FedNow system that will also offer 24/7 payment clearing services to retail customers is anticipated to be launched in 2023.

Payment systems are inextricably linked to the implementation of monetary policy—i.e., how the central bank sets overnight

*We thank Jason Allen, Michael Boutros, James Chapman, Jonathan Chiu, Chris Sutherland, and internal seminar participants at the Bank of Canada for helpful comments. All remaining errors are our own. Author contact: Bank of Canada, Ottawa, Ontario, K1A 0G9; E-mail: jwitmer@bankofcanada.ca.

interbank rates. The demand for reserves in a standard model of monetary policy implementation (e.g., Poole 1968) is generated by how uncertainty over interbank payment flows affects the use of central bank borrowing and lending facilities. 24/7 payment settlement has the potential to change both the nature of payment uncertainty as well as the use of central bank facilities. For example, demand for reserves could be a function of whether the central bank provides access to its lending facilities (i.e., intraday credit) only during standard operational hours or always. This paper aims to understand how demand for reserves is a function of those hours. If the central bank does not provide access to an after-hours central bank lending facility, a bank needs to have positive reserve balances greater than the payment amount to process a given payment. While a bank could establish a credit line to borrow from another bank to meet the payment, that other bank would also be worried about an inability to process payments. In either case, an extra dollar of reserves in the after-hours market provides a benefit in that it helps banks avoid having insufficient funds to process payments in the after-hours market. Does this matter for overnight interbank interest rates? Under what conditions will this matter, and is the impact different for different implementation frameworks? To answer these questions, we adjust the standard model of monetary policy implementation to incorporate after-hours payment shocks.

Traditional models (e.g., Poole 1968 and Bech and Keister 2013) have a payment shock that occurs while banks have access to central bank facilities. Thus, in these models, the cost and probability of accessing these facilities influences the interbank rate. If a payment shock occurs when banks do not have access to this facility, then banks need to factor the cost of having insufficient funds in the after-hours period into their marginal benefit of an extra dollar of reserves. Once we start thinking about 24/7 settlement, it opens up several questions related to monetary policy implementation, which this paper attempts to answer.[1]

[1]For example, in the United States, Fedwire Funds Service operates until 6:30 pm and the National Settlement Service (NSS) operates until 5:30 pm. The Federal Reserve is considering expanding their hours to 24/7 to provide a liquidity management tool to support a 24/7 real-time gross settlement (RTGS) service.

Some models consider how differential access to central bank facilities and segmentation in the overnight market affect the interbank interest rate (Bech and Klee 2011; Martin et al. 2013; Armenter and Lester 2015; Williamson 2019). In these papers, access to central bank facilities is segmented by participant. In our model, all participants have the same access, but that access is segmented by time. Like these other models, segmentation affects interbank interest rates. We also extend the baseline model to two periods, an intraday trading period and an after-hours trading period. In different contexts, other models also extend the baseline model to multiple periods (e.g., by looking at reserve averaging over multiple trading periods as in Ennis and Keister 2008).

We provide the conditions under which after-hours payments can have an effect on interbank interest rates. When after-hours payment volatility is material relative to intraday payment volatility, banks will have an increased demand for reserves. This increased demand increases with the volatility of the after-hours payment shock and is precautionary in that banks want to hold extra reserves to avoid having insufficient funds in the after-hours session. When the expected penalty cost is sufficiently large, banks will want to borrow more than the minimum requirement from the central bank. How this all affects interbank rates depends on the monetary policy implementation framework. Interbank rates in monetary policy frameworks that naturally have large reserves will not be affected much by such a change. On the other hand, there will be upward pressure on interbank rates in frameworks that typically have zero or low levels of reserves.

While the central bank can intervene by providing more aggregate reserves to offset this upward pressure, this could be more challenging if the volatility of the after-hours payment shock fluctuates. This could happen, for instance, if the after-hours period is longer in certain periods such as weekends or holidays. We therefore investigate how a change in the volatility of the after-hours payment shock affects the volatility of the overnight rate. When reserves are sufficiently large, changes in after-hours payment volatility do not matter. When reserves are smaller, changes in after-hours payment volatility result in volatility in the overnight interbank rate, absent a central bank response.

Finally, we examine the impact on the interbank overnight rate of having payments spread across two payment systems—a traditional intraday system and a 24/7 system. Parallel payment systems exist in several jurisdictions. When there is not real-time settlement between the two systems, we show that the overnight rate could be higher or lower than the overnight rate in one single payment system. In the extreme, when settlement across the two systems is completely restricted, the central bank must supply the appropriate level of reserves in each of the two systems if it wants to implement its target rate in both systems.

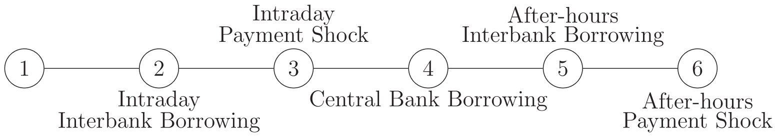
In practice, several central banks have already implemented payment systems with 24/7 retail payments, but overnight interbank rates still trade close to target in their jurisdictions. Our model would imply that either (i) uncertainty about retail payment flows in the after-hours session is small in these jurisdictions, or (ii) the level of reserves in these jurisdictions is large enough such that there is little chance that banks will have insufficient funds to process payment flows in the after-hours session. However, should more payment flows migrate to the 24/7 system, our model would suggest that this could lead to deviations from the target interest rate. Further, the implementation of a 24/7 payment system in countries that operate a system with low reserves or plan to return to such a framework could be very different than the experience thus far.

2. Model

2.1 Model Timing

Our model extends Bech and Keister (2013) and Boutros and Witmer (2019) and consists of six stages. We assume a continuum of perfectly competitive banks indexed by $i \in [0, 1]$. The first four stages presented in Figure 1 are standard in the literature. Banks borrow from (and lend to) each other during the day. After this borrowing and lending window is over, banks are subject to a payment shock. If they are short reserve balances after this payment shock, they must borrow from the central bank at rate r_X to make up the shortfall. If they have excess reserves, these get deposited with the central bank and earn interest r_R. We depart from the standard models by introducing after-hours trading in stage 5 and an after-hours

Figure 1. Model Timing



payment shock in stage 6. What distinguishes this after-hours payment shock from the intraday payment shock is that we assume that banks do not have recourse to the central bank borrowing facility in the after-hours market.

Banks begin the day in stage 1 with reserves, R^i, bond holdings, B^i, and deposit liabilities, D^i. Aggregate reserves are defined as $R = \int_i R^i \mathrm{d}i$. Bond holdings are exogenous and fixed throughout the day. We include bond holdings to be consistent with the literature (e.g., Bech and Keister 2013), but bonds are irrelevant to most of the analysis. Banks cannot choose the size of their deposits, the deposit rate (r_D) is fixed, and deposits only change when a bank experiences a payment shock.

In the standard model, a bank becomes a net lender ($\Delta^i_{intra} < 0$) or net borrower ($\Delta^i_{intra} > 0$) in stage 2 to position itself for the intraday payment shock it experiences in stage 3. In our model with after-hours payments shocks (i.e., shocks that happen after clearing and settlement of intraday balances), the banks' decisions are going to change. Specifically, a bank must also consider the effect of a penalty cost resulting from a large after-hours payment shock on its profitability. That is, it is not only minimizing the penal borrowing and lending rates associated with the central bank facilities, but is also minimizing penalty costs of having insufficient funds to process payments in the after-hours market.

In stage 3, after the trading session is closed, each bank experiences an intraday payment shock, $\epsilon^i_{intraday}$. This payment shock is independent and is identically and normally distributed with mean zero and standard deviation σ_G. We denote the cumulative distribution function of this shock $G(\epsilon^i_{intraday})$. This payment shock lowers the bank's reserves on the asset side of its balance sheet and correspondingly lowers its deposits on its liabilities side.

In stage 4, after the intraday payment shock, each bank borrows X_i from the central bank to meet or exceed its required level of

Table 1. Bank i's Stage 4 Balance Sheet

Assets	**Liabilities**
B^i Bonds $R^i + \Delta^i_{intra} - \varepsilon^i_{intraday} + X^i$ Reserves	$D^i - \varepsilon^i_{intraday}$ Deposits Δ^i_{intra} Interbank Borrowing X^i Central Bank Borrowing

reserves $K \geq 0$. We assume that each bank has the same required level of reserves, and that this required level of reserves is positive. The aggregate reserve requirement is defined as $K = \int_i K \mathrm{d}i$. We will focus our attention on frameworks with positive excess aggregate reserves (e.g., $R - K \geq 0$) that are commonly used in practice.

Assumption 1. *Aggregate reserves are greater than or equal to the aggregate reserve requirement.*

Banks that must borrow from the central bank do so at a rate of r_X. Each bank will borrow

$$X^i \geq \max\{0, K - (R^i + \Delta^i_{intra} - \epsilon^i_{intra})\}. \tag{1}$$

At a minimum, the bank must borrow at least enough to meet its minimum reserve requirement. We leave open the possibility that a bank may want to borrow more than this to help reduce the potential penalty cost of having insufficient funds in the after-hours session. At the end of stage 4, the bank earns r_K on its required reserves and $r_R < r_X$ on any reserves in excess of its required reserves. Table 1 illustrates bank i's balance sheet at the end of stage 4, before the after-hours payment shock.

In the stage 5 after-hours session, banks may be able to borrow from one another before the realization of the after-hours payment shock (similar to how they can borrow from each other during the day). We assume that banks can access interbank trading in the after-hours session with probability $1 - p$, $p \approx 0$ or $p = 1$. We make this assumption to see how the ability to trade in the after-hours market affects our model results. When $p = 1$, there is no after-hours trading session. When $p \approx 0$, the after-hours session resembles a session in which all banks can freely trade and lend from one another.

In this case, we assume $p \approx 0$ instead of $p = 0$ to generate a unique equilibrium when the penalty cost is high and banks may want to borrow more than the minimum from the central bank.

When they can access the after-hours market, a bank can be either a net lender ($\Delta^i_{after} < 0$) or a net borrower ($\Delta^i_{after} > 0$) in the after-hours market.

ASSUMPTION 2. *Each bank pays back both its intraday and after-hours loans at the beginning of the next day's session.*

In this sense, the intraday loan is an overnight loan, since it spans both before and after the central bank compensation of reserve balances in stage 4. The after-hours loan, in contrast, may be considered intraday because it is paid back before stage 4 the next day. The assumption about after-hours loans is for ease of exposition: when banks have access to an after-hours trading session ($p \approx 0$) we could instead assume banks make after-hours loans that are paid back after stage 4 the next day and obtain similar results.

In the final stage, each bank receives a payment shock in the after-hours market, ϵ^i_{after}. The after-hours payment shock is independent and is identically and normally distributed with mean zero and standard deviation σ_F. We denote the cumulative distribution function of this shock $F(\epsilon^i_F)$.

ASSUMPTION 3. *The volatility of the after-hours payment shock is smaller than the volatility of the intraday payment shock*: $\sigma_F \leq \sigma_G$.

This assumption reflects the fact that the after-hours time period is smaller than the intraday period. Also, most countries that offer 24/7 settlement do so for retail payment flows only, which would result in after-hours payment flows being smaller then intraday payment flows. Nonetheless, this assumption is inconsequential for most of the results in this paper, with the exception of Corollary 1.

The bank cannot meet its after-hours payment if

$$\epsilon^i_{after} \geq R^i + \Delta^i_{intra} + \Delta^i_{after} - \epsilon^i_{intra} + X^i. \tag{2}$$

If the bank cannot meet its payment, it suffers a penalty cost s on Z^i, each dollar of payment it is unable to make.

$$Z^i = \max\{0, \epsilon^i_{after} - (R^i + \Delta^i_{intra} + \Delta^i_{after} - \epsilon^i_{intra} + X^i)\} \tag{3}$$

ASSUMPTION 4. *Commercial banks can freely overdraft during the intraday session. During the after-hours session, in contrast, banks suffer a penalty cost, s, should they have insufficient funds to process a payment.*

This difference is a consequence of the central bank policy choice of not providing access to the central bank facilities in the after-hours period.[2] If access to central bank facilities and daylight credit was provided 24/7, our model would collapse to the standard model. We assume that the penalty cost of running out of funds during the day is zero. In reality there is a small cost to accessing intraday credit from the central bank (see Ennis and Weinberg 2007 for a model including intraday credit).[3] Our model reflects a situation where the after-hours penalty cost is materially larger than the small cost of intraday daylight credit.[4]

This penalty cost is modeled in reduced form as a negative profit if the bank is unable to make a payment. It could represent, for example, stigma associated with inability to make a payment, lost clients as a result of the inability to make a payment on their behalf, or late-payment charges embedded in contracts with clients.

2.2 Bank Behavior

Banks earn r_B on their bond holdings and pay r_D on their deposit holdings, both of which are exogenously determined. After the after-hours payment shock, bank i's realized profits are therefore

[2]The Federal Reserve is considering whether to provide intraday credit on a 24/7 basis with the implementation of its FedNow system. For more information, please refer to the following press release: https://www.federalreserve.gov/newsevents/pressreleases/files/other20190805a1.pdf.

[3]These costs include the opportunity cost of posting collateral with the central bank and any interest rates associated with daylight credit. Since 2011, the Federal Reserve assesses daylight overdraft charges on uncollateralized daylight overdrafts, which on average represent about 5 percent of daylight overdrafts. See https://www.federalreserve.gov/paymentsystems/psr_data.htm and https://www.federalreserve.gov/paymentsystems/files/psr_overview.pdf.

[4]We could adapt the model to include a non-zero intraday penalty cost if the bank's reserve position falls below zero in stage 4. This would not change the main insights from the model, so we assume that this penalty cost is zero during the day.

$$\begin{aligned}\pi^i &= r_B B^i - r_D(D^i - \epsilon^i_{intra}) + r_K K - r_{intra}\Delta^i_{intra} \\ &\quad - r_X X^i + r_R(R^i + X^i + \Delta^i_{intra} - \epsilon^i_{intra} - K) \\ &\quad - r_{after}\Delta^i_{after} - s * Z^i. \end{aligned} \tag{4}$$

We work backwards to find the bank's optimal behavior. In stage 5, when banks have access to the after-hours market they will choose their net interbank after-hours borrowing Δ^i_{after} to maximize their expected profits in the stage, $E_5[\pi^i_5]$ (i.e., expectations over the last two terms in the above equation). The numerical subscript on the expectations operator and profit variable indicates that these expectations are on bank profits as of stage 5. Banks will maximize:

$$\mathrm{E}_5[\pi^i_5] = -r_{after}\Delta^i_{after} - s\int_{\epsilon^i_Z}^{\infty}(\epsilon^i_{after} - \epsilon^i_Z)\mathrm{dF}(\epsilon^i_{after}). \tag{5}$$

In this equation, the threshold before the bank is expected to experience the penalty cost, $\epsilon^i_Z \equiv R^i + \Delta^i_{intra} + \Delta^i_{after} - \epsilon^i_{intra} + X_i$, is equal to the amount of reserves after after-hours trading is complete. This takes into account the amount of reserves at the beginning of the day, less the intraday payment shock, plus the amount of central bank borrowing after the intraday payment shock and the amount borrowed in the intraday and after-hours markets.

It will also be useful to define the same threshold in the absence of access to after-hours trading, $\epsilon^i_{Z,\Delta^0_{after}} \equiv R^i + \Delta^i_{intra} - \epsilon^i_{intra} + X_i$. This is the same as ϵ^i_Z, except that after-hours trading is set equal to zero. This is useful for examining the case where banks do not have the ability to trade in the after-hours session.

The value of Δ^i_{after} that maximizes the expected after-hours profits in Equation (5) is given by the following first-order condition:

$$r_{after} = s(1 - F(\epsilon^i_Z)). \tag{6}$$

Given that banks borrow and lend from each other at the same rate in the after-hours market (r_{after}), it follows from Equation (6) that banks will trade with each other such that they have the same $\epsilon^i_Z \equiv \epsilon_Z$. The bank's expected trading (in stage 4, before the after-hours payment shock) in the after-hours market can thus be written as

$$E_4[\Delta^i_{after}] = \epsilon_Z - (R^i + \Delta^i_{intra} - \epsilon^i_{intra} + X_i). \tag{7}$$

In stage 4, banks will take into account that they may be able to trade in the after-hours market, and will choose their central bank borrowing X_i to maximize the expected value of their profits in that stage, $E_4[\pi^i_4]$, subject to the constraint on central bank borrowing in Equation (1). They will maximize the Lagrangian:

$$\mathcal{L}^i = \mathrm{E}_4[\pi^i_4] + \lambda^i(X^i - \max\{0, K - (R^i + \Delta^i_{intra} - \epsilon^i_{intra})\}), \tag{8}$$

where

$$\begin{aligned}\mathrm{E}_4[\pi^i_4] = (r_R - r_X)X_i - (1-p)[r_{after}E_4[\Delta^i_{after}] \\ + s\int_{\epsilon_Z}^{\infty}(\epsilon^i_{after} - \epsilon_Z)\mathrm{dF}(\epsilon^i_{after})] \\ - ps\int_{\epsilon^i_{Z,\Delta^0_{after}}}^{\infty}(\epsilon^i_{after} - \epsilon^i_{Z,\Delta^0_{after}})\mathrm{dF}(\epsilon^i_{after}).\end{aligned} \tag{9}$$

This yields the following first-order condition and complementary slack condition with $\lambda_i \geq 0$:

$$r_X - r_R = (1-p)r_{after} + ps(1 - F(\epsilon^i_{Z,\Delta^0_{after}})) + \lambda_i \tag{10}$$

$$\lambda_i(X^i - \max\{0, K - (R^i + \Delta^i_{intra} - \epsilon^i_{intra})\}) = 0. \tag{11}$$

Because all banks face the same after-hours interbank rate, r_{after}, it follows from Equation (10) that if there are unconstrained banks (i.e., for whom $\lambda_i = 0$), they will borrow from the central bank until they have the same $\epsilon^i_{Z,\Delta^0_{after}} \equiv \epsilon^u_{Z,\Delta^0_{after}}$. We can define $\epsilon^u_{Z,\Delta^0_{after}}$ as the value of $\epsilon^i_{Z,\Delta^0_{after}}$ that solves Equation (10) with equality:

$$r_X - r_R = (1-p)r_{after} + ps(1 - F(\epsilon^u_{Z,\Delta^0_{after}})). \tag{12}$$

PROPOSITION 1. *No bank will borrow more (from the central bank) than the minimum required to meet its reserve requirement if*

$$r_X - r_R \geq (1-p)r_{after} + ps(1 - F(K)). \tag{13}$$

Proof. Because of the reserve requirement, all banks will be holding at least K in the after-hours session. Equations (10) and (13) imply that $\lambda_i > 0$ for all banks and hence all banks are constrained by the minimum reserve requirement. Intuitively, the marginal cost of borrowing from the central bank is greater than the marginal expected penalty cost in the after-hours session when all banks hold at least K, so no bank will want to borrow more than the minimum from the central bank. □

In this case, there will be no precautionary borrowing to alleviate the potential after-hours penalty cost. From Equations (12) and (13), it follows that all banks will be constrained if $K > \epsilon^u_{Z,\Delta^0_{after}}$.[5]

This means that banks will borrow from the central bank when they experience an intraday payment shock larger than the threshold ϵ^i_X:

$$\epsilon^i_{intra} \geq \epsilon^i_X \equiv R^i + \Delta^i_{intra} - max(\epsilon^u_{Z,\Delta^0_{after}}, K). \tag{14}$$

In stage 2, banks will take into account that they can borrow from the central bank and trade in the after-hours market. The expected amount of borrowing in the after-hours interbank market ($E_2[\Delta^i_{after}]$) is based on the expected borrowing from the central bank, which itself is a function of the bank's intraday trading (Δ^i_{intra}). However, borrowing funds in the intraday interbank market does not decrease after-hours borrowing one-for-one, since borrowing an extra dollar of reserves during the intraday session makes it less likely that the bank will need to borrow from the central bank at

[5]If $s < r_X - r_R$, there is no value of $\epsilon^u_{Z,\Delta^0_{after}}$ that will solve Equation (12), in which case we assume $K > \epsilon^u_{Z,\Delta^0_{after}}$. That is, no bank will want to borrow from the central bank and pay a higher cost, $r_x - r_R$, to avoid a probability of paying a lower cost s.

the end of the day. Taking expectations as of stage 2 of Equation (7) produces

$$E_2[\Delta^i_{after}] = \epsilon_Z - (R^i + \Delta^i_{intra} + \int_{\epsilon^i_X}^{\infty} (\epsilon^i_{intra} - \epsilon^i_X)\mathrm{dG}(\epsilon^i_{intra})). \tag{15}$$

By taking the derivative of Equation (15) with respect to Δ^i_{intra}, borrowing an additional dollar in the intraday market will reduce expected after-hours borrowing by $G(\epsilon^i_X)$.

Given this expected after-hours borrowing, banks will choose their net interbank intraday borrowing Δ^i_{intra} to maximize the expected value of their profits:

$$\begin{aligned} \mathrm{E}_2[\pi^i] = {} & r_B B^i - r_D D^i + r_K K^i - r_{intra}\Delta^i_{intra} \\ & + r_R\epsilon^i_X + (r_R - r_X)\int_{\epsilon^i_X}^{\infty} (\epsilon^i_{intra} - \epsilon^i_X)\mathrm{dG}(\epsilon^i_{intra}) \\ & - (1-p)\Big[r_{after}E_2[\Delta^i_{after}] + s\int_{\epsilon_Z}^{\infty} (\epsilon^i_{after} - \epsilon_Z)\mathrm{dF}(\epsilon^i_{after})\Big] \\ & - ps\Big[\int_{-\infty}^{\epsilon^i_X}\int_{\epsilon^i_{Z,\Delta^0_{after}}}^{\infty} (\epsilon^i_{after} - \epsilon^i_{Z,\Delta^0_{after}})\mathrm{dF}(\epsilon^i_{after})\mathrm{dG}(\epsilon^i_{intra}) \\ & + \int_{\epsilon^i_X}^{\infty}\int_{max(\epsilon^u_{Z,\Delta^0_{after}},K)}^{\infty} (\epsilon^i_{after} - max(\epsilon^u_{Z,\Delta^0_{after}}, K)) \\ & \mathrm{dF}(\epsilon^i_{after})\mathrm{dG}(\epsilon^i_{intra})\Big]. \end{aligned} \tag{16}$$

Relative to a standard Poole (1968) model, the bank's expected profit in (16) includes two groups of extra terms. The first group (with terms multiplied by $1-p$) concerns the bank's expected profit when it has access to the after-hours trading session. It includes a term that accounts for the expected profit from its expected borrowing and lending in the after-hours market. It includes another term that accounts for the penalty cost, s, of falling short of funds in the after-hours session when the bank is able to trade with other banks in stage 5. This term has the same threshold for each bank since their after-hours trading will make their after-hours reserves position the same. As the integral suggests, the bank only pays this cost if the after-hours payment shock, ϵ^i_{after}, exceeds the threshold ϵ_Z. This term, unlike the first term in this group, is not affected by the bank's intraday interbank trading.

The second group of terms (with terms multiplied by p) represent the penalty cost of falling short of funds in the after-hours session when the bank is unable to trade with other banks in stage 5. The first term in this group represents the expected penalty cost if the intraday payment shock is small enough such that the bank does not borrow from the central bank. The second term in this group represents the expected penalty cost if the bank does borrow from the central bank. Banks borrowing from the central bank will borrow such that they all have the same threshold before the penalty cost affects their profitability.

Formally, banks will choose Δ^i_{intra} to maximize their expected profits in Equation (16), resulting in the following first-order condition:

$$\begin{aligned} r_{intra} &= r_R + (r_X - r_R)(1 - G(\epsilon^i_X)) + (1-p)r_{after}G(\epsilon^i_X) \\ &\quad + ps\int_{-\infty}^{\epsilon^i_X}[1 - F(\epsilon^i_{Z,\Delta^0_{after}})]\mathrm{dG}(\epsilon^i_{intra}). \end{aligned} \tag{17}$$

2.3 Equilibrium

2.3.1 Equilibrium Overnight Rate

DEFINITION. *An equilibrium consists of interest rates r_{intra} and r_{after} and individual bank net borrowing decisions (Δ^i_{intra}) and (Δ^i_{after}) such that*

(i) Banks choose Δ^i_{after} to maximize expected profits in the stage 5 after-hours session, as in (5).

(ii) Banks choose X^i to maximize expected profit when borrowing from the central bank in stage 4, as in (9).

(iii) Banks choose Δ^i_{intra} to maximize expected profit in stage 2, as in (16).

(iv) The interbank markets are closed systems that clear, that is, $\Delta_{intra} = \int_i \Delta^i_{intra} \mathrm{d}i = 0$ *and* $\Delta_{after} = \int_i \Delta^i_{after} \mathrm{d}i = 0$*. Moreover, for banks that do not have access to the after-hours trading session (*$i \in$ *no access),*

$$\int_{i\ \in\ no\ \ access} \Delta^i_{intra} \mathrm{d}i = \int_{i\ \in\ access} \Delta^i_{intra} \mathrm{d}i = \Delta_{intra} = 0.$$

By the first-order condition in (17), and by the fact that banks are all subject to the same reserve requirement K, they will have the same ϵ^i_X in equilibrium. That is, since r_{intra} is the same for all banks, there is only one value of ϵ^i_X that will solve (17).

By market clearing $\Delta_{intra} = \Delta_{after} = 0$, and it follows that

$$\epsilon_X = \int_i \epsilon^i_X \mathrm{d}i = R - max(\epsilon^u_{Z,\Delta^0_{after}}, K) \tag{18}$$

$$\epsilon_Z = \int_i \epsilon^i_Z \mathrm{d}i = R + \int_{\epsilon_X}^{\infty} (\epsilon^i_{intra} - \epsilon_X) \mathrm{dG}(\epsilon^i_{intra}). \tag{19}$$

Given that each bank trades to hold the same threshold amounts before the intraday payment shock, it follows from the equilibrium definition that $\epsilon_X = \epsilon^i_X$ and $\epsilon_Z = \epsilon^i_Z$.

There are two equilibrium possibilities, depending on the size of the penalty cost. In the first equilibrium, the penalty cost is small enough such that all banks borrow just the minimum from the central bank to meet their reserve requirement. This will be the case if $K \geq \epsilon^u_{Z,\Delta^0_{after}}$. In this case, the cost of borrowing from the central bank will be more than the expected penalty cost s in the after-hours session when the bank holds the required level of reserves (e.g., Equation (13)). In this low penalty cost equilibrium, the threshold amount in (18) is the same as in the standard Poole model (i.e., $\epsilon_X = R - K$).

In the second equilibrium, the penalty cost is sufficiently large such that some banks borrow more than the minimum from the central bank ($\epsilon^u_{Z,\Delta^0_{after}} > K$). In this high penalty cost equilibrium, the threshold for central bank borrowing (ϵ_X) in Equation (18) is lower than in the low penalty cost equilibrium, since banks will want to borrow more from the central bank to reduce the penalty cost in the after-hours session. And, since they borrow more from the central bank and thus will have more funds in the after-hours session, it follows that the threshold to experience the penalty cost ϵ_Z as in Equation (19) will be higher in the high penalty cost equilibrium.

The first-order condition in Equation (17) can now be written as a function of these aggregate threshold amounts, which themselves depend on the aggregate bank's balance sheet (i.e., as in Equations (18) and (19)).

$$\begin{aligned} r_{intra} = {} & (r_R + (1-p)r_{after})G(\epsilon_X) + r_X(1 - G(\epsilon_X)) \\ & + ps \int_{-\infty}^{\epsilon_X} [1 - F(R - \epsilon^i_{intra})] \mathrm{dG}(\epsilon^i_{intra}) \end{aligned} \tag{20}$$

2.3.2 Comparison with Overnight Rate in the Absence of an After-Hours Payment Shock

Our second proposition compares the overnight rate in Equation (20) with the overnight rate in a standard model, $r_{Poole} \equiv r_R G(R-K) + r_X(1 - G(R-K))$.

PROPOSITION 2. *The overnight rate in the presence of an after-hours payment shock will be weakly greater than the overnight rate in the absence of one:*

$$\begin{aligned} r_{intra} = {} & r_{Poole} + (r_X - r_R)[G(R-K) - G(\epsilon_X)] \\ & + (1-p)r_{after}G(\epsilon_X) \\ & + ps \int_{-\infty}^{\epsilon_X} [1 - F(R - \epsilon^i_{intra})] \mathrm{dG}(\epsilon^i_{intra}). \end{aligned} \tag{21}$$

Equation (21), which substitutes the Poole rate into Equation (20), shows that the overnight rate with an after-hours payment shock is equal to the overnight rate in a standard model, r_{Poole}, plus three additional positive terms to account for the benefit of additional funds in the after-hours session in avoiding the expected penalty cost in the after-hours session. This holds in both the high penalty cost and low penalty cost equilibrium. The first additional term, $(r_X - r_R)[G(R-K) - G(\epsilon_X)]$, represents borrowing beyond the minimum reserve requirement and will put upward pressure on the overnight rate relative to the Poole model. In the low penalty cost equilibrium, this term is equal to zero. In the high penalty cost equilibrium, this term is positive and the overnight rate with a high penalty cost will be above the Poole rate. The second additional term accounts for the expected penalty cost when there is an after-hours trading session, while the third additional term accounts for the cost if there is no after-hours trading session.

Just like in the standard Poole model, the overnight rate in the intraday session will still be bounded by the central bank deposit and lending rates, r_R and r_X, in the presence of a penalty cost in the after-hours session.

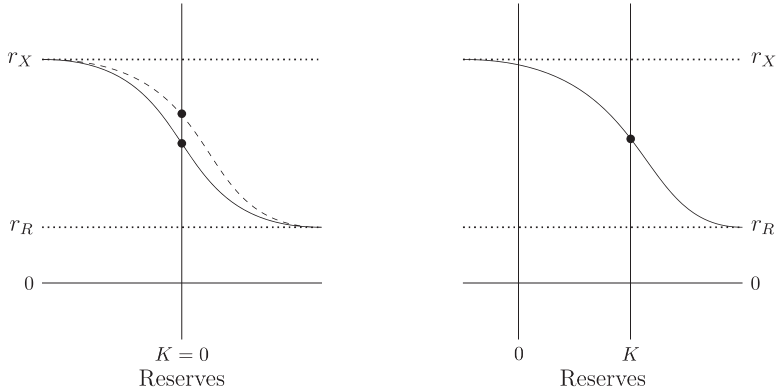
$$r_R \leq r_{intra} \leq r_X \tag{22}$$

It is easy to see that the overnight rate will be bounded from below by the central bank deposit rate. The standard Poole rate is bounded from below by the central bank deposit rate, and Proposition 2 shows that the overnight rate is higher than the Poole rate in the presence of an after-hours penalty cost. To see that it is bounded from above by the central bank lending rate, substitute Equation (10) into the first-order condition in Equation (20). This yields $r_{intra} = r_X - \int_{-\infty}^{\epsilon_X} \lambda^i \leq r_X$.

Figure 2 illustrates how the demand for reserves changes in the presence of an after-hours payment shock. In the graph on the left-hand side, where there are no required reserves, the demand for reserves in the presence of an after-hours payment shock (dashed line) is higher than the demand for reserves in the absence of this shock (solid line). This will, if anything, put upward pressure on the overnight interbank rate.

Interestingly, the demand for reserves is unaffected if aggregate reserves are very large or very small. When aggregate reserves are very large, there is almost zero probability that a bank would experience an after-hours payment shock that fully drains its reserves, so the expected penalty cost associated with having insufficient reserves is negligible. On the other hand, when aggregate reserves are very small (large, negative value), banks will almost surely borrow from the central bank at the end of the day. Therefore, trading away an additional dollar in the interbank market will have no effect on the probability of having insufficient funds in the after-hours period, since the bank would borrow an extra dollar from the central bank at

Figure 2. Demand for Overnight Reserves



Note: The left-side graph illustrates the case where the required level of reserves, K, is equal to zero. The solid line in this graph illustrates the demand for reserves when there is no after-hours payment shock (i.e., the traditional Poole model). The dashed line represents the demand for reserves when there is an after-hours payment shock. The dots represent the equilibrium allocation and rates when the level of reserves is also equal to zero, showing that the interbank rate could trade above the middle of the corridor when there is an after-hours payment shock. In the right-side graph, the required level of reserves, K, is a large positive number. In this graph, the demand for reserves is unaffected by the presence of an after-hours payment shock and the dashed line and the solid line coincide. The dot represents the equilibrium allocation and rate when the level of reserves is equal to the required level of reserves, showing that, with large required reserves, the interbank rate could still trade in the middle of the corridor when there is an after-hours payment shock.

the end of the day before the after-hours payment shock. As such, its interbank trading would have no effect on its reserve position before experiencing the after-hours payment shock.

2.3.3 Impact of After-Hours Payment Shock in Different Monetary Policy Implementation Frameworks

The following corollaries state the conditions under which the overnight rate will equal the Poole rate under different central bank operating frameworks. For these corollaries, we will focus on the equilibrium with a small penalty cost, since the high penalty cost equilibrium produces an overnight rate above the Poole rate. Specifically, we examine three different monetary policy implementation frameworks commonly used in practice, defined as follows:

(i) *A floor framework*: In this framework, reserves are sufficiently large such that the overnight rate in the standard Poole model (i.e., no after-hours payment shock) is equal to the deposit rate (i.e., $r_{Poole} = r_R$). For this to be the case, $G(R-K) \approx 1$.

(ii) *A zero-reserve corridor framework*: $R = K = 0$. The overnight rate in the standard Poole model will be equal to the midpoint of the central bank lending and deposit rates (i.e., $r_{Poole} = \frac{r_R + r_X}{2}$).

(iii) *A positive reserve corridor framework*: $R = K > 0$. The overnight rate in the standard Poole model will also be equal to the midpoint of the central bank lending and deposit rates (i.e., $r_{Poole} = \frac{r_R + r_X}{2}$).

Corollary 1. *In a floor framework with a low penalty cost ($K \geq \epsilon^u_{Z,\Delta^0_{after}}$), the overnight rate is equal to the Poole interest rate when*

- *participants can almost surely trade in the after-hours session ($p \approx 0$), or*
- *excess reserves are larger than a threshold value A: $R-K \geq A$, with $G(A) \approx 1$ and $F(R-A) \approx 1$.*

Proof. In the low penalty cost equilibrium, Equation (21) simplifies to

$$r_{intra} = r_{Poole} + (1-p)r_{after}G(\epsilon_X) + ps\int_{-\infty}^{\epsilon_X}[1 - F(R - \epsilon^i_{intra})]\mathrm{dG}(\epsilon^i_{intra}). \tag{23}$$

$F(\epsilon_Z) \geq F(R-K)$ and, by Assumption 3 and the definition of a floor, $F(R-K) \geq G(R-K) \approx 1$. Therefore, $F(\epsilon_Z) \approx 1$ and from Equation (6) $r_{after} = 0$, meaning the first additional term in the above equation is equal to zero. The second additional term in the above equation is equal to zero when either $p \approx 0$ by the first condition in this corrollary or $F(K) \approx 1$ by the second condition, leaving $r_{intra} = r_{Poole}$. $\square$

The result is intuitive. In a floor framework, the expected value of the penalty cost is near zero since there is close to zero probability that the after-hours payment shock will reduce the bank's reserves below zero.

Visually, this can be illustrated with the central bank supplying a large quantity of reserves in Figure 2 (left-side graph), such that the supply of reserves is a vertical line that intersects with the demand for reserves at r_R. At this point, the demand for reserves is the same as it would be in the absence of an after-hours payment shock.

COROLLARY 2. *In a zero-reserve corridor framework* ($R = K = 0$) *with a low penalty cost, the overnight rate will equal the Poole rate only when*

- *the volatility of the overnight payment shock is relatively small (i.e., when* F *and* G *are normally distributed and* $\Phi(\frac{\sigma G}{\sigma_F}\phi(0)) \approx 1$), *and*
- *participants can almost surely trade in the after-hours session* ($p \approx 0$).

Proof. Substituting $R = K = 0$ and $p \approx 0$ into Equation (23) and writing in terms of standard normal distributions yields the result, $r_{intra} = r_{Poole} + \frac{s}{2}(1 - \Phi(\frac{\sigma G}{\sigma_F}\phi(0)))$. Since $\Phi(\frac{\sigma G}{\sigma_F}\phi(0)) \approx 1$), $r_{intra} = r_{Poole}$. $\square$

Intuitively, when the volatility of the intraday payment shock increases (holding σ_F constant), there is more borrowing from the central bank in stage 4. In particular, when G is a normal distribution, aggregate borrowing from the central bank increases linearly with the volatility of the intraday payment shock in a zero-reserve corridor. This aggregate borrowing increases the aggregate supply of reserves in the after-hours market. When this aggregate borrowing becomes sufficiently large and banks can trade in the after-hours market, the expected penalty cost of the after-hours payment shock for each bank is approximately 0. This means that $r_{after} = 0$ and thus $r_{intra} = r_{Poole}$.

Determining the optimal level of aggregate reserves to target the Poole rate in a zero-reserves corridor is more challenging in the presence of material after-hours payment shocks. In the absence of an after-hours payment shock, the central bank simply needs to target aggregate reserves equal to the aggregate reserve requirement, $R = K$. With a material after-hours payment shock, the central bank needs to understand the demand for reserves to determine the amount of aggregate reserves to supply to the market. This will depend on, among other things, the size of the penalty cost s and the magnitude of after-hours payment shocks, σ_F. This can be seen in Figure 2 (left-side graph). With $R = K = 0$, the equilibrium interest rate (the intersection of the supply and the dashed demand curve) will be higher than the equilibrium interest rate in the absence of an overnight payment shock (the intersection of the supply and the solid demand curve).

An alternative for the central bank could be to establish a higher required reserves amount. This leads to our next corollary.

COROLLARY 3. *In a positive-reserve requirement corridor system ($R = K > 0$), the overnight rate will equal the Poole rate when the aggregate reserve requirement is sufficiently large (e.g., $F(K) \approx 1$).*

Proof. Substituting $R = K$ and $F(K) \approx 1$ into Equation (6) shows that $r_{after} = 0$ and thus the first additional term in Equation (23) equals zero. The second additional term in Equation (23) is also equal to zero when $F(K) \approx 1$, which means $r_{intra} = r_{Poole}$. □

The intuition for this result is the same as the intuition for Corollary 1. When K is sufficiently large, all banks will hold enough reserves such that there is a near zero probability that the after-hours payment shock will bring the bank's level of reserves to zero, where it will begin to experience the penalty cost. The higher amount of required reserves shifts the demand curve for reserves to the right, as seen in Figure 2 (right-side graph). Thus, the demand for reserves is the same as in the standard case.

2.4 Effect of After-Hours Payment Shock Volatility on the Equilibrium

A good monetary policy implementation framework should be characterized by low volatility in the overnight rate (e.g., Bindseil 2016).

In our model, we can examine how the volatility of the overnight rate is affected by the volatility of σ_F, the volatility of the after-hours payment shock. After-hours payment volatility could fluctuate, for instance, if the length of the after-hours session fluctuates (i.e., over a weekend).

To see the impact on day-to-day volatility in the overnight rate, we now assume that the after-hours payment shock volatility is a random variable that is symmetrically distributed around its mean value. We label this mean value as $\overline{\sigma_F}$. Before stage 1 each day, all banks see the realization of the payment shock volatility random variable. Since the random variable is realized before trading begins, it will not change the model's results on a given day but will change results from one day to the next. The overnight rate is also now a random variable because it is a function of the after-hours payment volatility, which is itself a random variable.

The volatility of the overnight rate with respect to changes in the volatility of the after-hours payment shock can thus be expressed as

$$\sigma_{r_{intra}} = \sqrt{E[(r_{intra}(\sigma_F) - E[r_{intra}(\sigma_F)])^2]}. \tag{24}$$

Next, we can take a Taylor-series expansion of the overnight rate around the mean value of the after-hours payment shock volatility:

$$r_{intra}(\sigma_F) \approx r_{intra}(\overline{\sigma_F}) + \frac{\partial r_{intra}}{\partial \sigma_F}(\sigma_F - \overline{\sigma_F}). \tag{25}$$

This and the symmetric distribution of σ_F means we can approximate the expected overnight rate by $E[r_{intra}(\sigma_F)] \approx r_{intra}(\overline{\sigma_F})$. By substituting this and Equation (25) into Equation (24), the effect of volatility of the after-hours payment shock volatility on overnight rate volatility can be approximated as

$$\sigma_{r_{intra}} \approx \left|\frac{\partial r_{intra}}{\partial \sigma_F}\right| \sqrt{E[(\sigma_F - \overline{\sigma_F})^2]}. \tag{26}$$

This expression suggests that transmission of volatility of after-hours payment shock volatility ($\sigma_{\sigma_F} \equiv \sqrt{E[(\sigma_F - \overline{\sigma_F})^2]}$) to volatility of the overnight rate will depend on $\frac{\partial r_{intra}}{\partial \sigma_F}$. All else equal, an overnight rate that is more responsive to changes in after-hours payment shock volatility will be more volatile.

Some central banks could adjust aggregate reserves to limit the impact of changes in after-hours payment shock volatility on overnight rate volatility. In this case, aggregate reserves would also be a function of after-hours payment shock volatility (e.g., $R = R(\sigma_F)$). That is, the central bank may offset the effect of changes in after-hours payment shock volatility by changing aggregate reserves. Given this, we can take the derivative of the overnight rate with respect to after-hours payment volatility and substitute this derivative into Equation (26) to write overnight rate volatility as

$$\sigma_{r_{intra}} \approx \left| s\Phi\left(\frac{\epsilon_X}{\sigma_G}\right)\phi\left(\frac{\epsilon_Z}{\sigma_F}\right)\frac{\epsilon_Z}{\sigma_F^2} + \frac{\partial r_{intra}}{\partial R}\frac{\partial R}{\partial \sigma_F}\right| \sigma_{\sigma_F}. \tag{27}$$

In the absence of a central bank response ($\frac{\partial R}{\partial \sigma_F} = 0$), changes in after-hours payment volatility will weakly increase volatility of the overnight rate. Two sufficient circumstances where changes in after-hours payment volatility do not affect overnight rate volatility are as follows:

- A floor framework. In Corollary 1 we showed that $F(\epsilon_Z) \approx 1$ in a floor framework. When $F(\epsilon_Z) \approx 1$, $\phi(\frac{\epsilon_Z}{\sigma_F}) \approx 0$ and hence $\sigma_{r_{intra}} \approx 0$.
- A positive-reserve requirement corridor system, where the aggregate reserve requirement is sufficiently large (e.g., $F(K) \approx 1$). Similarly, when $F(K) \approx 1$, $\phi(\frac{\epsilon_Z}{\sigma_F}) \approx 0$ and hence $\sigma_{r_{intra}} \approx 0$.

In both of these cases, volatility of after-hours payment volatility has no effect on the overnight rate. In this case reserves are sufficiently large such that a change in after-hours payment volatility does not change the probability of running out of funds in the after-hours session.

In a zero-reserve requirement corridor system, on the other hand, volatility in after-hours payment volatility will increase the volatility of the overnight rate. While the central bank could offset this impact on overnight rate volatility by adjusting aggregate reserves in responses to changes in after-hours payment volatility, it may be difficult for the central bank to determine the necessary adjustment. The required adjustment requires knowledge of the demand for overnight funds (i.e., $\frac{\partial r_{intra}}{\partial R}$) as well as knowledge about the volatility of the after-hours payment shock and intraday payment shock.

3. Model Applications

3.1 Optimal Reserve Requirements in a Corridor System

The previous subsection suggests that the central bank can effectively maintain a corridor system if it chooses required reserves that are sufficiently high. With a high reserve requirement, the expected penalty cost of having insufficient funds in the after-hours market is negligible, and volatility of after-hours payment shock volatility does not affect the overnight rate.

However, a large positive-reserve requirement corridor system may not necessarily be optimal if the central bank has other concerns beyond the overnight rate setting. To illustrate this, we assume that a bank now faces a balance sheet cost that is an increasing function of balance sheet size, $c(D) \geq 0$, $c'(D) \geq 0$, and $c''(D) \geq 0$. We focus on the scenario where banks can trade in the after-hours session ($p \approx 0$) and banks are in the low penalty cost equilibrium. Bank bond holdings remain exogenous.[6] Given the balance sheet identity and the fact that $R = K$ in a corridor, aggregate deposits are equal to $B + K$. Then, suppose that the central bank chooses K to minimize costs to the banking system:

$$\text{Cost} = c(B+K) + s\int_{\epsilon_Z^i}^{\infty} (\epsilon_{after}^i - \epsilon_Z)\mathrm{dF}(\epsilon_{after}^i). \tag{28}$$

Then, it is simple to see that the central bank will choose a reserve requirement such that

$$c'(B+K) = s\left(1 - \Phi\left(\frac{K + \sigma_G\phi(0)}{\sigma_F}\right)\right). \tag{29}$$

This illustrates the trade-off where increasing reserve requirements reduces the expected after-hours penalty cost but increases

[6]This section is only meant to illustrate that there are other considerations beyond controlling the overnight rate. A general equilibrium analysis of social welfare where banks also choose their bond holdings is beyond the scope of the paper, but we point the interested reader to Martin et al. (2013), Canzonari, Cumby, and Diba (2017), and Williamson (2019).

the balance sheet cost of banks. Assuming it can't set negative reserve requirements, a central bank would prefer a zero-reserve requirement corridor to a positive-reserve requirement corridor if the marginal balance sheet cost is large relative to the penalty cost at $K = 0$:

$$c'(B) \geq s\left(1 - \Phi\left(\frac{\sigma_G \phi(0)}{\sigma_F}\right)\right). \tag{30}$$

3.1.1 Return on Required Reserves in a Corridor

If this social optimum requires a large reserve requirement, banks may see this large reserve requirement as a tax (e.g., see the discussion in Lipscomb, Martin, and Wiggins 2017). However, by adjusting the return on required reserves, we show how the central bank could make its choice of reserve requirement coincide with that which banks would choose themselves.[7]

We assume banks choose their required relative reserves before the start of intraday trading, under the assumption that the central bank will supply aggregate reserves equal to the aggregate reserve requirement ($R = K$). This is a slight departure from Baughman and Carapella (2018), given that in their model voluntary reserve targets adjust to central bank's supply of reserves rather than the other way around. In our model, the introduction of a balance sheet cost also allows us to avoid the problem of banks setting infinite targets. In the absence of a balance sheet cost, banks could make infinite profits if $r_D < r_K$ by increasing their reserve requirements and this would cause reserve targets to converge to $+\infty$.

The aggregate bank will choose K to maximize expected profits:

$$\begin{aligned} \mathrm{E}[\pi] = {} & r_B B - r_D(B + K) - c(B + K) + r_K K \\ & + (r_R - r_X)\int_0^\infty \epsilon^i_{intraday} \mathrm{dG}(\epsilon^i_{intraday}) \\ & - s\int_{\epsilon^i_Z}^\infty (\epsilon^i_{after} - \epsilon_Z)\mathrm{dF}(\epsilon^i_{after}). \end{aligned} \tag{31}$$

[7] Baughman and Carapella (2018) develop a model of voluntary reserve targets and show the potential advantages of such a model over other models.

The first-order condition is

$$r_D + c'(B+K) = r_K + s\left(1 - \Phi\left(\frac{K + \sigma_G \phi(0)}{\sigma_F}\right)\right). \tag{32}$$

The left-hand side of the first-order condition represents the marginal cost of an additional dollar of required reserves. This is the bank's marginal funding cost, and it includes the deposit rate, as well as the cost associated with increasing the bank's balance sheet. The right-hand side of the equation is the marginal benefit of an additional dollar of required reserves. It consists of two components. The first is r_K, the rate at which the central bank compensates required reserves. The second is the reduction in the expected penalty costs associated with having insufficient funds to process after-hours transactions.

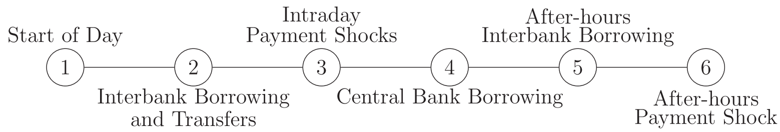
In this illustrative example, if the central bank could set the rate on required reserves equal to the deposit rate ($r_D = r_K$), the social optimum level of required reserves would coincide with that which the banks would choose themselves. In a more general equilibrium setup, r_D would not be fixed and it may be more challenging for the central bank to set $r_D = r_K$. Nonetheless, the main point of this subsection should still hold more generally: a central bank can adjust r_K so that its choice of reserve requirement coincides with what banks would choose themselves in the optimum.

3.2 Multiple Payment Systems

In this section, we analyze how the results are affected by the operation of two interlinked payment systems: one which operates only during the day (labeled a traditional system) and has access to the central bank borrowing and lending facilities during operating hours, and one which operates 24/7 and has access to the central bank deposit facility but not the lending facility (outside of operating hours).

Several jurisdictions currently have multiple payment systems, with one payment system providing 24/7 access. The 24/7 system is usually a system for retail payment flows, and there are only a few cases where 24/7 retail payment flows are accommodated within the large-value payment system or are settled in real time

Figure 3. Model Timing with Two Payment Systems



in the large-value payment system.[8] Tompkins and Olivares (2016) show that some jurisdictions, however, have infrequent settlement between the 24/7 retail payment system and the large-value payment system, with the ability to pre-fund the 24/7 system when the large-value payment system is closed. How does the presence of multiple payment systems affect monetary policy implementation? How will parameters such as the frequency of settlement between two interoperating systems affect the overnight rate?

To analyze this setup, we model infrequent settlement as a restriction on the ability to transfer between the two systems after trading is complete in stage 2. Figure 3 illustrates how the setup is modified. We focus on the low cost equilibrium and assume that all banks can trade in the after-hours session. For ease of notation, we assume that banks begin the day with zero reserves in the 24/7 payment system. This does not affect the results since, in stage 2, in addition to borrowing in the interbank market, we assume banks can also transfer funds between the two payment systems. We denote the net transfer from the traditional payment system to the 24/7 payment system $T^i \geq 0$. We interpret this transfer between systems as both a settlement between the systems as well as the ability to pre-fund the 24/7 system.

[8]See Tompkins and Olivares (2016), Figure 13 for a list of jurisdictions with 24/7 retail payment systems and non-24/7 wholesale payment systems, and the interoperability between these systems. Moreover, in 2016, the Committee on Payments and Market Infrastructures of the Bank for International Settlements listed 11 countries with retail payment systems that operated on a near-24/7 basis, plus an additional 8 countries that had plans to implement one (Committee on Payments and Market Infrastructures 2016). Canada, for instance, plans to introduce a 24/7 retail payments system that will operate alongside its large-value payment system, which operates with traditional hours (Payments Canada 2017).

In stage 3, after the trading session is closed, each bank experiences an intraday payment shock, $\epsilon^i_{intraday}$, in the traditional payment system. This is the same payment shock as in the baseline setup, with cumulative distribution function $G(\epsilon^i_{intraday})$. For simplicity and comparison with the earlier results, we assume there is no intraday payment shock in the 24/7 payment system.

After experiencing the intraday payment shock, as before, banks borrow from the central bank in stage 4 if they are in a negative excess reserve position in the traditional system.[9] Then, the banks earn r_R on positive balances they hold in either system. They also pay r_X on their borrowing from the central bank.

In stage 5, banks will still choose their net interbank after-hours borrowing Δ^i_{after} to maximize their expected profits in the after-hours market, given the payment shock it is exposed to in stage 6 in the 24/7 payment system:

$$\mathrm{E}_5[\pi^i_5] = -r_{after}\Delta^i_{after} - s\int_{\epsilon^i_{Z,T}}^{\infty}(\epsilon^i_{after} - \epsilon^i_{Z,T})\mathrm{dF}(\epsilon^i_{after}), \quad (33)$$

where the threshold, $\epsilon^i_{Z,T} \equiv T^i + \Delta^i_{after}$, is a little different because it accounts for the effect of transfers between the two systems. Since central bank borrowing only affects the reserve position in the traditional system, it does not affect the threshold for experiencing the penalty cost in the 24/7 system.

Like before, the first-order condition from Equation (33) provides the after-hours interbank rate that maximizes the expected after-hours profits:

$$r_{after} = s(1 - F(\epsilon^i_{Z,T})). \quad (34)$$

Equation (34) implies that banks will trade with each other until they have the same $\epsilon^i_{Z,T} \equiv \epsilon_{Z,T}$.

[9]If we assume that banks can transfer between systems in stage 4, that reserve requirements apply to balances in the 24/7 system, and that there is a reserve requirement of 0 in the traditional system at the end of the day, results will be identical to the baseline setup of a single system. Each bank will make inter-system transfers in stage 4 such that there will be a zero balance in the traditional payment system.

Each bank will choose their borrowing, lending, and transfer activity in stage 2 to maximize expected profits:

$$\begin{aligned} \mathrm{E}[\pi^i] = {} & r_B B^i - r_D D^i + r_K K^i - r_{intra}\Delta^i_{intra} \\ & + r_R(\epsilon^i_X - T^i) + (r_R - r_X) \\ & \int_{\epsilon^i_X - T^i}^{\infty} (\epsilon^i_{intraday} - (\epsilon^i_X - T^i))\mathrm{dG}(\epsilon^i_{intraday}) \\ & - r_{after}(T^i + \Delta^i_{after}) - s\int_{\epsilon_{Z,T}}^{\infty} (\epsilon^i_{after} - \epsilon_{Z,T})\mathrm{dF}(\epsilon^i_{after}). \end{aligned} \tag{35}$$

Maximizing expected profits produces two first-order conditions. After combining these two conditions and aggregating across all banks, the optimality conditions can be written as

$$r_{intra} = r_R + (r_X - r_R)\left(1 - \Phi\left(\frac{R - K - T}{\sigma_G}\right)\right)$$

$$r_{intra} = r_R + s\left(1 - \Phi\left(\frac{T}{\sigma_F}\right)\right).$$

The first equation represents the marginal value of reserves in the traditional system, and the second equation represents the marginal value of reserves in the 24/7 system. Since both equations have the overnight interbank rate on the left-hand side, it suggests that banks will transfer funds between the two systems until the marginal value of reserves is equal across both systems.

The marginal value of reserves suggested by the first equation is a small departure from the standard model. Transfers to the 24/7 system will reduce reserves in the traditional system and increase the probability that the bank will have to borrow from the central bank. Hence, it will put upward pressure on the overnight rate relative to the Poole rate.

The marginal value of reserves in the 24/7 system is a function of two factors. First, banks will earn interest on reserves on funds held in the 24/7 system. Second, an extra dollar of reserves in the 24/7 system lowers the likelihood that the bank will be short of funds in the after-hours session, thus reducing the expected penalty cost of being short of funds.

Overall, the interbank rate in the two-system environment may be higher or lower than the interbank rate in a single system. On the one hand, banks will borrow more in the traditional system than they do in a single system, since transfers to the 24/7 system reduce reserves in the traditional system and hence increase the amount of interbank borrowing. Also, some banks will end up with positive balances at the end of the day in the traditional system that cannot be used to reduce the probability of experiencing the penalty cost in the 24/7 system (since post-intraday shock transfers are not allowed). This will put upward pressure on the interbank rate in a dual system, relative to a single system. On the other hand, in a single system, the marginal benefit of an additional dollar of reserves includes both the marginal benefit of reducing the expected cost of central bank borrowing, as well as the marginal benefit of reducing the expected penalty cost. Because it contains both of these marginal benefits, and the interbank rate reflects these marginal benefits, the interbank rate in a single system could be higher. Depending on which of these two competing effects dominates, the rate could be higher in a single system or in a dual system.

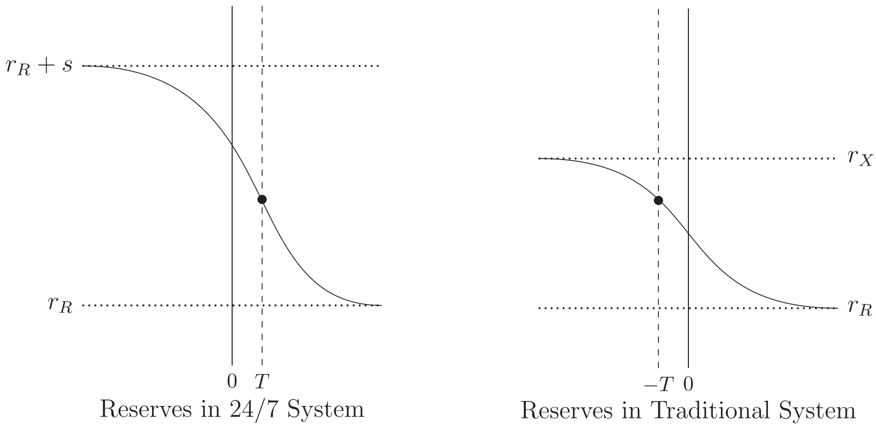
Figure 4 illustrates how these transfers affect the overnight interbank rate. When both the traditional system and the 24/7 system begin the day with zero reserves, the marginal value of funds in the 24/7 system (Figure 4, left-side graph) is greater than the marginal value of funds in the intraday system (Figure 4, right-side graph). Since the marginal value of funds is higher in the 24/7 system, participants will have incentives to move reserves into that system. They continue doing so until the marginal value of funds in the two systems is equal.

3.2.1 Tighter Restrictions on Transfers

We assumed that transfers can occur during the interbank trading period. An even more extreme assumption would be to completely restrict transfers across the two systems. This could represent, for example, those 24/7 retail payment systems that settle in the large-value payment system only once a day (Tompkins and Olivares 2016).

A complete restriction on transfers generates segmented markets across the two systems. Because transfers cannot take place in

Figure 4. Overnight Rate in Multiple Systems



Note: The graphs represent overnight trading in the two systems when the required level of reserves, K, and aggregate level of reserves, R, are both equal to zero. The left-side graph illustrates the equilibrium allocation and rate in the 24/7 system, while the right-side graph illustrates the equilibrium allocation and rates in the traditional system. Assuming that both systems begin the day with zero aggregate reserves, T represents the transfers between the two systems that occur in stage 2. Specifically, transfers occur until the overnight rates in the two different systems are equal.

stage 2 when trading occurs, the marginal value of funds in the two systems will (most likely) be different from each other. Specifically, there will be an overnight interbank rate for the traditional system ($r_{intra,traditional}$) and an overnight interbank rate for the 24/7 system ($r_{intra,24/7}$). Reserves in the two systems are now denoted $R_{traditional}$ and $R_{24/7}$, respectively. Thus, in stage 2, banks are either trading funds in the traditional system with each other, or trading funds in the 24/7 system with each other:

$$r_{intra,traditional} = r_R + (r_X - r_R)\left(1 - \Phi\left(\frac{R_{traditional} - K}{\sigma_G}\right)\right)$$

$$r_{intra,24/7} = r_R + s\left(1 - \Phi\left(\frac{R_{24/7}}{\sigma_F}\right)\right).$$

Restricting transfers completely insulates the traditional trading market from effects of the after-hours payment shock: the interbank rate in the traditional system is equal to the Poole rate. To implement the same rate in the 24/7 system, the central bank would need to determine the appropriate supply of reserves in the 24/7 system based on how these reserves affect $r_{intra,24/7}$ in the formula above.

4. Conclusion

Our paper shows how changes in the payment system could have implications for overnight interest rates. Specifically, monetary policy implementation frameworks that naturally have a large amount of reserves are less affected by a move to 24/7 payment settlement in our model. More broadly, while our model focuses on the effect of 24/7 payment settlement on interbank rates, it can also be applied to other factors that increase the benefits of reserves. For example, the penalty costs of having insufficient funds after hours could be interpreted as a cost of having insufficient reserves relative to a target level of reserves that could be driven by regulation or other factors.

Further, our model is derived in a centralized market, so it does not say anything about trading volumes or dispersion of traded rates. While we believe our model delivers the important implications of 24/7 settlement, a search model with a decentralized market (e.g., Afonso and Lagos 2015) could provide additional implications for trading activity.

Finally, we assume that the intraday shock process and other underlying features of the system, such as the number and characteristics of participants, are invariant to payment system modernization. In practice, these features may adjust to a lengthening of the payment period. We leave a more detailed exploration of these drivers to future work.

References

Afonso, G., and R. Lagos. 2015. "Trade Dynamics in the Market for Federal Funds." *Econometrica* 83 (1): 263–313.

Armenter, R., and B. Lester. 2015. "Excess Reserves and Monetary Policy Normalization." Working Paper No. 16-33, Federal Reserve Bank of Philadelphia.

Baughman, G., and F. Carapella. 2018. "Voluntary Reserve Targets." Finance and Economics Discussion Series No. 2018-032, Board of Governors of the Federal Reserve System.

Bech, M., and T. Keister. 2013. "Liquidity Regulation and the Implementation of Monetary Policy." Working Paper No. 432, Bank for International Settlements.

Bech, M., and E. Klee. 2011. "The Mechanics of a Graceful Exit: Interest on Reserves and Segmentation in the Federal Funds Market." *Journal of Monetary Economics* 58 (5): 415–31.

Bindseil, U. 2016. "Evaluating Monetary Policy Operational Frameworks." Remarks at the 2016 Economic Policy Symposium sponsored by the Federal Reserve Bank of Kansas City, held at Jackson Hole, Wyoming, August 25–27.

Boutros, M., and J. Witmer. 2019. "Monetary Policy Implementation in a Negative Rate Environment." *Journal of Money, Credit and Banking* 52 (2–3): 441–70.

Canzonari, M., R. Cumby, and B. Diba. 2017. "Should the Federal Reserve Pay Competitive Interest on Reserves?" *Journal of Money, Credit and Banking* 49 (4): 663–93.

Committee on Payments and Market Infrastructures. 2016. "Fast Payments — Enhancing the Speed and Availability of Retail Payments." Technical Report, Bank for International Settlements.

Ennis, H. M., and T. Keister. 2008. "Understanding Monetary Policy Implementation." *Economic Quarterly* (Federal Reserve Bank of Richmond) 94 (3, Summer): 235–63.

Ennis, H. M., and J. A. Weinberg. 2007. "Interest on Reserves and Daylight Credit." *Economic Quarterly* (Federal Reserve Bank of Richmond) 93 (2, Spring): 111–42.

Lipscomb, L., A. Martin, and H. Wiggins. 2017. "Why Pay Interest on Required Reserve Balances." Blog Entry, Liberty Street Economics Blog, Federal Reserve Bank of New York, September 25.

Martin, A., J. McAndrews, A. Palida, and D. Skeie. 2013. "Federal Reserve Tools for Managing Rates and Reserves." Staff Report No. 642, Federal Reserve Bank of New York.

Norman, B., R. Shaw, and G. Speight. 2011. "The History of Interbank Settlement Arrangements: Exploring Central Banks' Role in the Payment System." Working Paper No. 412, Bank of England.

Payments Canada. 2017. "Modernization Target State." Technical Report.

Poole, W. 1968. "Commercial Bank Reserve Management in a Stochastic Model: Implications for Monetary Policy." *Journal of Finance* 23 (5): 1177–95.

Tompkins, M., and A. Olivares. 2016. "Clearing and Settlement Systems from Around the World: A Qualitative Analysis." Staff Discussion Paper No. 2016-14, Bank of Canada.

Williamson, S. D. 2019. "Interest on Reserves, Interbank Lending, and Monetary Policy." *Journal of Monetary Economics* 101 (January): 14–30.

Tracing the Impact of the ECB's Asset Purchase Program on the Yield Curve*

Fabian Eser,[a] Wolfgang Lemke,[a] Ken Nyholm,[a]
Sören Radde,[b] and Andreea Liliana Vladu[a]
[a]European Central Bank
[b]Point72

We trace the impact of the European Central Bank's (ECB) asset purchase program (APP) on the yield curve. Exploiting granular information on sectoral asset holdings and ECB asset purchases, we construct a novel measure of the "free-float of duration risk" borne by arbitrageurs. We include this supply variable in an arbitrage-free term structure model in which central bank purchases reduce the free-float of duration risk and hence compress term premia. We estimate the stock of current and expected future APP holdings to reduce the 10-year yield by almost 1 percentage point. This reduction is persistent, with a half-life of five years.

JEL Codes: C5, E43, E52, E58, G12.

*We are grateful for comments from two anonymous referees and to Massimo Rostagno for helpful discussions. We also thank Ulrich Bindseil, Alexander Düring, Robin Greenwood, Natacha Valla, and Dimitri Vayanos, as well as participants at the CFE-CM Statistics Conference in Pisa, University of Zürich Finance Seminar, Frankfurt Goethe University Macro Seminar, University of York Asset Pricing Workshop, Bank of England seminar, European Finance Association, German Economic Association, and ECB internal seminars for useful comments. Christopher Greiner and Alexia Ventula Veghazy provided excellent research assistance. Sören Radde worked on this paper while employed by the ECB, prior to joining Point72. The information, views, and opinions expressed herein are solely his own and do not necessarily represent the views of Point72 or its affiliates. Point72 and its affiliates are not responsible for, and did not verify for accuracy, any of the information contained herein. ECB disclaimer: "The views expressed are those of the authors and do not necessarily reflect those of the European Central Bank." Corresponding author (Wolfgang Lemke) e-mail: wolfgang.lemke@ecb.europa.eu.

1. Introduction

We trace the impact of the European Central Bank's (ECB) asset purchase program (APP) on the euro area sovereign yield curve at announcement and over time. The ECB launched the APP in January 2015 by pledging the purchase of €60 billion of public and private sector securities a month from March 2015 until at least September 2016, amounting to €1.1 trillion. Successive rounds of recalibrations of the APP in December 2015, March 2016, December 2016, October 2017, and June 2018 took the eventual size of the portfolio to around €2.6 trillion by the end of net purchases in December 2018, corresponding to around 25 percent of euro area GDP. The ECB thus joined other major central banks, such as the Federal Reserve, in employing large-scale purchases—also known as "quantitative easing (QE)"—to provide monetary policy accommodation in the proximity of the effective lower bound by seeking to lower longer-term yields.[1]

For our analysis we deploy an affine term structure model in which central bank asset holdings compress term premia by reducing the amount of duration risk borne by arbitrageurs, building on Li and Wei (2013). In affine term structure models that are commonly used to study bond yield dynamics,[2] supply effects of securities do not play an explicit role. By contrast, the microfounded model by Vayanos and Vila (2021), featuring preferred-habitat investors and arbitrageurs, links the term premium to the amount of duration risk to be absorbed by the arbitrageurs: lower aggregate duration risk increases the risk-bearing capacity of the arbitrageurs, thereby decreasing risk compensation per unit of risk exposure (i.e., the "price of risk") and hence the term premium. It is the overall amount of duration risk that matters for the term premium. Therefore, a change in bond supply at a specific maturity affects not only that

[1]While the ECB previously embarked on outright asset purchases in the form of the Securities Markets Program (SMP), these purchases were categorically different from QE-type large-scale asset purchases and instead consisted of sterilized temporary interventions to provide liquidity to selected debt markets; see Eser and Schwaab (2016), Ghysels et al. (2016), and De Pooter, Martin, and Pruitt (2018).

[2]Cf. Kim and Wright (2005), Joslin, Singleton, and Zhu (2011), and Adrian, Crump, and Moench (2013).

maturity bracket but also term premia along the entire curve. Moreover, the model by Vayanos and Vila (2021) predicts that it is the whole sequence of current and discounted future aggregate duration in the market that determines current bond prices.

This link between bond supply and the term premium is captured in our term structure model by including a quantitative measure of duration risk in addition to standard level and slope yield curve factors. This allows us to study the term premium effect of the ECB's APP, which decreases the overall duration risk to be borne by arbitrageurs. Finally, as in Li and Wei (2013), we restrict the supply factor to not affect current and expected future short-term interest rates, thereby excluding a "rate signaling" channel of central bank asset purchases.[3]

Our empirical measure of duration risk in the market is inspired by the theory developed by Vayanos and Vila (2021). Rather than considering the exposure of all private investors, as in Li and Wei (2013) and Ihrig et al. (2018), we exploit security-level information on sectoral bond holdings from the ECB's Securities Holding Statistics (SHS) to develop a more granular measure. From total bond holdings we exclude not only bond holdings by domestic central banks and governments but also the portfolios of domestic hold-to-maturity investors as well as the foreign official sector, since these investor groups are unlikely to respond to changes in the supply and maturity structure of outstanding bonds. As a residual, we obtain the bond holdings of arbitrageurs. We weight these holdings

[3]Signaling effects of the ECB's non-standard monetary policy measures have been analyzed by, e.g., Andrade et al. (2016), Arrata and Nguyen (2017), Lemke and Werner (2020), and Altavilla, Carboni, and Motto (2021). Such effects have been found to be small in magnitude compared with the effects of the duration extraction channel. By contrast, based on a shadow-rate term structure model estimated for the overnight index swap (OIS) yield curve, Geiger and Schupp (2018) find that unconventional monetary policy shocks have a stronger impact on expected short rates than on the forward term premium up to seven years. We do not separately identify the role of reserves creation for term premium compression as Christensen and Krogstrup (2019) do based on Swiss data, as reserve- and supply-induced effects cannot be independently identified for QE programs involving purchases of long-term securities. We also abstract from flow effects of purchases, which are of a more temporary nature; see D'Amico and King (2013) and Kandrac and Schlusche (2013) for the United States, Joyce and Tong (2012) for the United Kingdom, as well as Schlepper et al. (2017) and De Santis and Holm-Hadulla (2020) for the euro area.

according to their duration and normalize them by the total duration supply of outstanding government bonds. We refer to the share of duration risk exposure borne by arbitrageurs relative to total duration risk supply as the "free-float of duration risk."

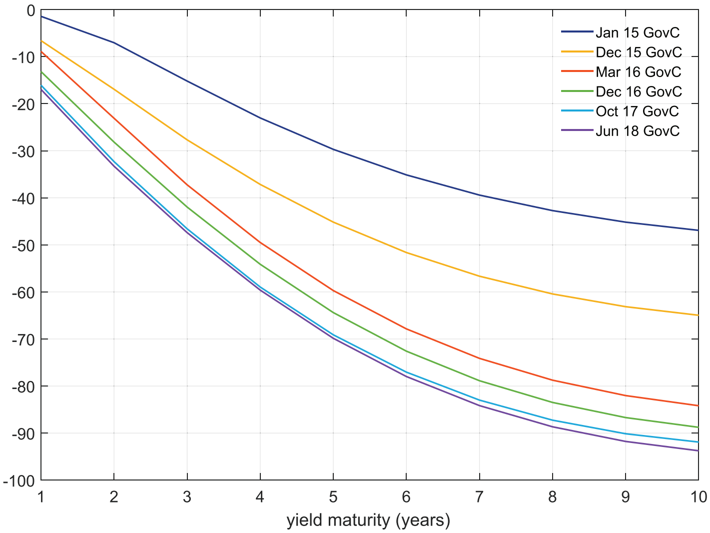
We estimate the model by minimizing the weighted sum of two fitting criteria. The first criterion measures the time-series fit of euro area sovereign bond yields (zero coupon, averaged across the largest four countries) over the period before markets started pricing large-scale asset purchases by the ECB. The second criterion is based on the fit of the cumulative yield decline over events (ECB press conferences and speeches) in the run-up to and around the announcement of the APP, which were perceived by markets to contain information on the forthcoming purchase program. We rely on this novel approach as our sample is relatively short: we can only construct our free-float measure from December 2009 based on the SHS data. Moreover, Eurosystem[4] bond holdings only became a significant source of variation in the free-float with the start of APP. This contrasts the U.S. experience, where the Federal Reserve's monetary policy portfolio exhibited significant variation already before the inception of its large-scale asset purchases (LSAPs).

We report four main results. First, we find that the APP has flattened the yield curve and compressed term sovereign premia considerably; see Figure 1. Specifically, at its onset in January 2015, the APP compressed 10-year sovereign term premia by around 50 basis points (bps), the impact has increased gradually as the APP has been expanded in length and volume, and the impact has reached around 95 bps in June 2018. The 5–95 percent confidence interval, which accounts for parameter uncertainty, ranges from 65 to 130 bps.

Second, we find that the term premia compression due to the APP is persistent. Based on the path of APP net purchases envisaged by the Governing Council in June 2018, and assuming a horizon for full reinvestments of 3 years, we estimate a half-life of around 5 years for the 10-year term premium impact. The fading of the impact over time reflects, to some extent, the aging of the portfolio, i.e., the gradual loss of duration as the securities held in the portfolio mature,

[4]The Eurosystem comprises the ECB and the 19 national central banks of the euro area member states.

Figure 1. Impact of Different APP Vintages on the Sovereign Yield Curve



Note: The figure shows the contemporaneous impact of the APP on the term structure of interest rates via the duration extraction channel. For the indicated dates t and maturities n, the respective point on the line provides an estimate of how much the sovereign n-period yield at the respective time t is compressed due to the impact on the term premium via the duration extraction channel.

as well as, in particular, the run-down of the portfolio that market participants anticipate to follow the reinvestment phase.

Third, the expected length of the reinvestment period after net purchases has a significant impact on term premia. The longer the reinvestment horizon, the larger the term premium impact. For example, under the counterfactual of no reinvestment, relative to an assumed reinvestment horizon of three years, the long-term interest rate would have been around 15 bps higher in June 2018.

Fourth, we use our model to make real-time predictions of the yield curve effect of the various APP recalibrations and compare these predictions with the observed yield curve reactions upon announcement, controlling for the expectations of APP parameters prevailing ahead of the announcement. We find that the model accounts well out of sample for the observed yield curve changes around APP announcements that implied major surprises regarding the future free-float.

Overall, our quantitative results for the yield impact of the APP (almost 50 bps in early 2015 and around 95 bps by mid-2018 for the 10-year big-four sovereign yield) are within the wide range of estimates reported in the literature. In comparing results, it is important to account for the different empirical approaches, as well as data and sample choices.[5] Several papers deploy event-study approaches, which focus (by design) on the surprise element of QE. While our econometric model also uses event information for the estimation, our impact estimates also exploit information on the full expected trajectory of ECB bond holdings. Quantifying the APP impact on euro area GDP-weighted average sovereign yields based on an event study, Altavilla, Carboni, and Motto (2021) find an impact of the initial APP package on the 10-year yield of 65 bps, while De Santis (2020) reports an APP impact on the 10-year yield of 72 bps over the period August 2014–October 2015 using a panel error correction specification with key APP dates informed by Bloomberg news stories. Bulligan and Delle Monache (2018) also conduct event-study regressions and find that news about the APP during September 2014 to July 2017 reduced the 10-year yield by around 50 bps. Andrade et al. (2016) conduct an event study over yields of individual sovereign bonds eligible under the APP and quantify the impact of the duration extraction channel of the initial APP package in early 2015 on bonds with a duration larger or equal to 10 years at 47 bps. Combining event information with survey-based evidence of investors' purchase expectations, Rostagno et al. (2021) estimate that APP has compressed the 10-year euro area sovereign yield by around 120 bps at the end of 2018. Blattner and Joyce (2020) consider an alternative GDP-scaled free-float measure and estimate, using a Bayesian vector autoregression (VAR), the impact of the

[5]Koijen et al. (2021) estimate a demand system for government bonds by instrumental variables, finding that portfolio rebalancing actions of euro area investors between March 2015 and December 2017 reduced the average yield of the sovereign debt outstanding of each of the four largest euro area countries between 40 and 60 bps. As they average across maturities, their figures correspond roughly to our seven-year yield impact of 85 bps. Our findings are also consistent with the overview paper of Hartmann and Smets (2018), who report that in December 2016 the cumulated and joint impact of the APP together with credit-easing measures and interest rate cuts amounted to around 150 bps for the euro area 10-year sovereign yield.

initial APP package on the euro area 10-year bond yield at around 30 bps.

Our estimates are also broadly in line with those obtained for the U.S. Federal Reserve purchase programs, despite differences in market environment and purchase modalities on the two sides of the Atlantic. Applying the model by Li and Wei (2013) to the United States, Ihrig et al. (2018) estimate a peak cumulative impact of the Federal Reserve's LSAPs and its Maturity Extension Program of around 125 bps for a purchase volume of $4.5 trillion. A direct comparison of our euro area peak impact estimates—around 95 bps in June 2018, with uncertainty bands ranging from 65 to 130 bps—with the U.S. figures is challenging due to factors such as a different sovereign bond market structure, a different global financial environment at the time of purchases, and a different allocation of purchases over time. Moreover, ideally a comparison would be based on a granular free-float measure of the type we construct, but its U.S. counterpart is not available to us.[6] Taking the size of the economy as a very rough yardstick of comparison, overall purchase volumes amount to around one-quarter of GDP in both cases. Hence, under such scaling, the U.S. impact would range in the upper part of our confidence band obtained for the APP.

The remainder of the paper is structured as follows. Section 2 explains the construction of our free-float measure and the yield data. Section 3 describes the model and inspects the mechanism of how central bank purchases affect the term premium. Section 4 outlines the estimation approach and documents the model fit. Section 5 reports the main results, i.e., the impact on the yield curve at different points in time, the persistence of those effects, the role of reinvestment, and the impact of selected APP recalibrations. It also sheds some light on the robustness of results. The last section concludes.

2. APP Duration Extraction and Euro Area Yields

We construct a theory-consistent measure of the free-float of duration risk, which enters our term structure model as supply variable

[6]Using GDP as a scaling variable, as in Li and Wei (2013) and Ihrig et al. (2018), instead of total bond supply (as in our baseline specification) leaves our estimates largely unchanged, as discussed in more detail in Section 5.4.

(Section 2.1); we explain how to project it into the future using official ECB communication and survey information (Section 2.2); and we introduce the yield curve data (Section 2.3). Our analysis focuses on the government debt and average yields of the four largest euro area countries (Germany, France, Italy, and Spain; henceforth "big four"). These countries accounted for around 80 percent of euro area sovereign debt and around 76 percent of euro area GDP at the end of 2016.

2.1 A Theory-Consistent Measure of the Free-Float of Duration Risk

As in Vayanos and Vila (2021), the term premium is affected by the amount of duration risk to be absorbed by arbitrageurs.[7] Motivated by their theoretical framework, we construct an empirical measure of the free-float of duration risk as follows:

$$\text{free-float of duration risk} = \frac{\text{duration-weighted bond holdings of arbitrageurs}}{\text{duration-weighted total bond supply}}. \quad (1)$$

The three key dimensions of this empirical measure of the free-float of duration risk are, first, which type of investors to count as arbitrageurs; second, the normalization of the free-float of duration risk with the total bond supply; and third, the range of securities considered for the measurement of duration risk.

Regarding the first dimension, we compute the duration-weighted bond holdings of arbitrageurs (the numerator in Equation (1)) in order to account for the role of different types of investors in the transmission of central bank asset purchases.[8] We divide

[7]See also Hamilton and Wu (2012). The appendix provides a more detailed discussion of the mapping between Vayanos and Vila (2021) and our model.

[8]To account for the duration of the bonds held by arbitrageurs, we consider the sectoral holdings in terms of their 10-year equivalents. The 10-year equivalent portfolio is a hypothetical portfolio that consists only of 10-year zero-coupon bonds and that has the same duration risk as the actual portfolio. The nominal amount (par value) of an individual bond j is converted into the 10-year equivalent using the following formula: $10y\ equivalent_j = nominal_j \cdot \frac{duration_j}{10}$. We use the maturity as a duration proxy. This measure has the advantage of

holding sectors into two groups—arbitrageurs and preferred-habitat investors—in line with Vayanos and Vila (2021). To distinguish the bond holdings of these two groups of investors, we exploit the granular information available in the Eurosystem Securities Holdings Statistics (SHS) on the sectors holding general government debt securities. The SHS data are available from 2009:Q4 at quarterly frequency. At the security level, these data provide information on the nominal value, the residual maturity, and the holding sectors. For euro area holdings, the SHS distinguishes the following holding sectors: monetary and financial institutions (MFI), money market funds (MMF), non-MMF investment funds, insurance corporations and pension funds (ICPF), other financial institutions, non-financial corporations (NFC), and households. By contrast, for foreign, i.e., non-euro area, holdings only a distinction between official and non-official portfolios is available.[9] In addition, we use information on

---

abstracting from endogenous feedback effects from yield levels on other portfolio duration measures, such as modified duration. For a portfolio with weighted average maturity $WAM = \sum_{j=1}^{bonds}(nominal_j \cdot maturity_j)/\sum_{j=1}^{bonds} nominal_j$, the 10-year equivalent of the portfolio is obtained by weighting the portfolio's nominal value by $\frac{WAM}{10}$.

[9]The information on foreign holdings in the SHS is subject to two reporting biases, which we address as follows. First, nominal holdings by foreign private investors are inflated due to a custodial over-reporting bias of foreign non-official holdings. The SHS holdings for investors outside the euro area are collected from custodians. Custodians are financial institutions that hold securities on behalf of their customers. However, the custodians may not always know the final investor, especially if the customers of the custodians are institutions transacting on behalf of a third-party customer. If these third-party customers are located in the euro area, then the holdings reported by custodians as foreign holdings are in fact domestic holdings. As a result, the sum of all domestic and foreign holdings of some securities as reported in the SHS can exceed the outstanding amount of these securities as reported by the ECB's Government Finance Statistics (GFS). To address the over-reporting bias coming from custodians, we benchmark the nominal value of total outstanding government bonds for each country as reported in the SHS against the corresponding information from the GFS. We then adjust the foreign sector holdings obtained from the SHS downwards so that the sum of outstanding amounts across all sectors from the SHS data matches the totals from the GFS. Second, during the preliminary SHS data collection period from 2009:Q4 to 2013:Q3 foreign official sector holdings were largely unreported, as the submission of these data was voluntary. To address this issue, we backcast the nominal foreign official sector holdings for all countries from their 2013:Q4 levels based on the dynamics of euro area external financial liabilities as reported in the International Monetary Fund's Coordinated Portfolio Investment Survey (CPIS)

Eurosystem holdings derived from the ECB-internal security-level data on sovereign bond purchases.

Our group of preferred-habitat investors comprises the official sector, both euro area and foreign, and on the private sector side ICPFs. The inclusion of official holdings in the preferred-habitat category reflects the fact that these tend to have narrowly described mandates, which limit the degree to which these type of investors can engage in arbitrage. The official holdings comprise foreign exchange reserves by non-euro area central banks, holdings of the intra-euro area general government sector, as well as Eurosystem portfolios. The latter include both monetary-policy-related sovereign bond holdings, such as those accumulated under the Securities Markets Program (SMP) and the Public Sector Purchase Program (PSPP), as well as holdings which are unrelated to monetary policy and subject to the Agreement on Net Financial Assets (ANFA). Similarly, out of the private sector investors we include ICPFs in the preferred-habitat group, as these tend to follow hold-to-maturity strategies, matching long-dated liabilities with long-dated assets, and they are subject to regulatory requirements. ICPFs are, thus, unlikely or limited in their ability to rebalance away from their preferred habitats.

Our group of arbitrageurs is made up of all the private sectors other than ICPFs. These include—for the euro area—MFIs (excluding the Eurosystem), MMFs, non-MMF investment funds, NFCs, and households. In terms of magnitude, in the group of arbitrageurs MFIs are by far the dominant private domestic-holding sector of euro area sovereign bonds. In addition, we include the foreign non-official sector in the group of arbitrageurs.

We present some summary statistics of the asset holdings of the different sectors in Table 1. Before the start of the APP, euro area MFIs held the largest portion of big-four sovereign bonds, followed by the official sector other than the Eurosystem, which mainly reflects foreign reserve holdings. On balance, close to 55 percent of all outstanding big-four government bonds were in the hands of arbitrageurs pre-APP. Differences in the pre-APP average maturity

and the Currency Composition of Official Foreign Exchange Reserves (COFER). In addition, we assume that the weighted average maturity (WAM) of foreign official sector holdings is constant over the preliminary SHS data collection period at the level of the average WAM of the pre-APP official reporting period.

Table 1. Sovereign Bond Holdings by Investor Type and Sector

	Holdings (€bn)		WAM (Years)	
	Pre-APP	**2018:Q2**	**Pre-APP**	**2018:Q2**
Arbitrageurs	3,281	2,632	6.1	6.7
MFI	1,334	1,141	5.0	5.4
Other Domestic	1,126	976	6.5	7.8
Foreign Non-official	815	514	7.3	7.7
Preferred-Habitat Investors	2,716	3,860	6.6	7.4
ICPF	1,009	1,241	10.8	10.6
Other Official	1,305	1,096	4.1	4.3
Eurosystem	402	1,523	4.5	7.0

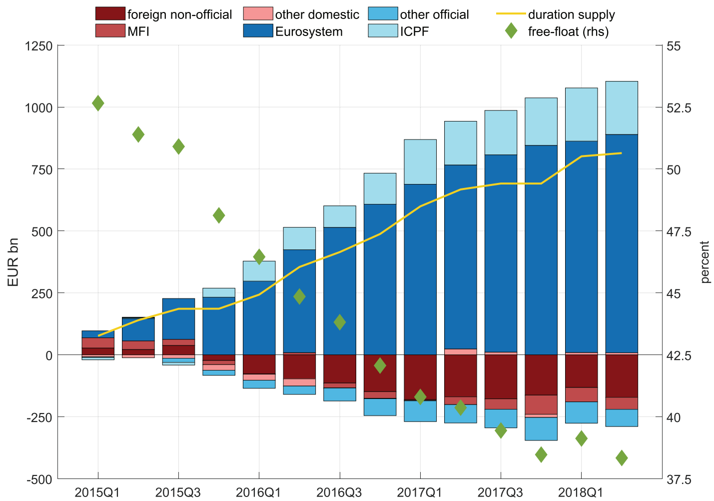
Note: The table reports the nominal value and WAM of sector-specific holdings of bonds issued by the general governments of France, Germany, Italy, and Spain. The pre-APP period refers to average sector holdings in the period 2013:Q4 to 2014:Q4. The sector "Other Official" includes domestic euro area governments and the foreign official sector; "Other Domestic" includes NFCs, households, and financial institutions other than banks.

of sectoral portfolios point to different investment strategies. For instance, MFI holdings tended to be concentrated in shorter maturity segments compared with the maturity distribution of all outstanding government bonds, while ICPFs held substantially longer-dated paper.

Since the start of the APP, a notable portfolio rebalancing has taken place across sectors.[10] The share of Eurosystem holdings in total outstanding big-four government bonds has risen from less than 7 percent to around 23 percent by mid-2018; see Table 1. ICPFs were the only investors who increased their holdings alongside the Eurosystem. All sectors classified as arbitrageurs, and among these most prominently banks and foreign non-official investors, have been net sellers of government bonds. As a result, the share of holdings by arbitrageurs as a fraction of total outstanding big-four government bonds has fallen from 55 percent to around 41 percent.

[10] See Bua and Dunne (2019) for a more detailed inspection deploying micro panel data from the investment fund industry domiciled and reporting in Ireland.

Figure 2. Evolution of Duration Supply and Its Absorption by Investor Type



Note: On the left axis, the chart shows cumulative changes in 10-year equivalent holdings of big-four general government bonds in €billion compared with pre-APP holdings as of 2014:Q4. On the right axis, the evolution of the free-float as a percentage of total duration supply is shown.

Figure 2 illustrates that since the start of the APP the Eurosystem has broadly offset the increase in the 10-year equivalent bond supply to the market through the issuance of big-four general government debt securities. ICPFs have increased their exposure to duration risk over the same period, which is consistent with their classification as preferred-habitat investors. By contrast, arbitrageurs, such as foreign private investors and, to a lesser extent, euro area banks, have reduced their relative exposure to duration risk since the start of the APP, as is evident from the material decline of the free-float.

Our classification of investor types as preferred habitat versus arbitrageurs is in line with the evidence on sectoral portfolio rebalancing in response to the PSPP in Koijen et al. (2017, 2021) and Bergant, Fidora, and Schmitz (2020). In particular, Koijen et al. (2017) and Koijen et al. (2021) identify the ICPF sector as the only one that increased its holdings euro area sovereign bonds—trading in the same direction as the ECB. The authors see the inelastic or even upward-sloping demand by ICPFs reflecting their need to match their long-dated liabilities, which implies a preferred habitat for ICPF investments. By contrast, in particular the foreign sector and banks, but also mutual funds and households, decreased their holdings. Regarding the inclusion of MMFs, NFCs, and households in our group of arbitrageurs, we acknowledge that these groups may not arbitrage across assets in the same way that MFIs do. However, we see the fact that MMFs, NFCs, and households have rebalanced their portfolios in the same direction as MFIs and the foreign private sectors, as supporting the classification of MMFs, NFCs, and households as arbitrageurs. In any case, the holdings of MMFs, NFCs, and households are small in relative terms.

By way of comparison, most studies in the literature either take all government bonds in the hands of the private sector as the numerator of the relevant supply variable (in Equation (1)), or they account for foreign official bond holdings to some extent, while the preferred-habitat behavior of ICPFs is not accounted for. Specifically, in the U.S. context Li and Wei (2013) consider U.S. Treasury securities in the hands of all private investors as the numerator of the bond supply measure. Similarly, D'Amico et al. (2012) also work with the privately held Treasury supply. By contrast, foreign official holdings are accounted for by Hamilton and Wu (2012) who use, inter alia, the average maturity of outstanding debt, where the weights are given by the share of non-official-sector debt holdings in total debt of the respective maturity, whereas Kaminska and Zinna (2020) rely on a weighted average of the total Treasury bond supply, bond supply held by the Federal Reserve, and foreign bond holdings. In the European context, Blattner and Joyce (2020) deploy a measure of maturity-weighted debt, from which they subtract a proxy of bond holdings by the foreign official sector. Compared with the literature, we believe that by excluding both the official sector holdings and the holdings of ICPFs we obtain a better approximation of the holdings of "arbitrageurs" along the lines of Vayanos and Vila (2021).

Turning to the second dimension—the normalization of the duration-weighted bond holdings of arbitrageurs—we normalize by the total supply of duration risk, i.e., the 10-year equivalent value of the nominal amount of outstanding government bonds of the big-four euro area jurisdictions; see the denominator in Equation (1).

We refer to the share of duration risk exposure of arbitrageurs in the total duration risk supply as the free-float of duration risk; see the green diamonds in Figure 2. For the estimation of the model the quarterly free-float series is interpolated linearly to obtain observations at monthly frequency.

Whereas in the theoretical setting of Vayanos and Vila (2021) the quantities of bonds enter directly as market values, some scaling is necessary in an empirical context, for a growing bond market would in principle lead term premia (which are a function of bond supply) to increase without bound. Moreover, scaling can also be motivated by the fact that any bond-holding variable measured in nominal level terms is likely to be persistent and thus challenging for econometric inference; see also Kaminska and Zinna (2020).[11]

Normalizing by total bond supply captures investors' need to bear more risk when debt supply increases. At the same time, choosing the total bond supply may not fully account for investors' capacity to bear more risk in a growing economy. This may be captured to some extent by scaling by GDP, which is the approach taken by Li and Wei (2013), Greenwood and Vayanos (2014), and Blattner and Joyce (2020). In any case, in our robustness section (see Section 5.4), we investigate the sensitivity of our empirical estimates to the choice of normalizing variable and also present results obtained by using GDP as a scaling variable. We find that the results obtained using GDP as a scaling variable are very similar in terms of magnitude. This reflects the fact that the evolution of the total debt supply and GDP are correlated. Moreover, the similarity of the results obtained by using total bond supply and GDP as scaling values suggests that the impact estimates are mainly driven by the change of the free-float induced by central bank purchases, which dominate the slower-moving variations in the scaling variable.

With regard to the third dimension—the range of securities considered for our measure of the free-float of duration risk—we focus on the general government bonds of the four largest euro area countries ("big four") purchased under the PSPP part of the APP. General government bonds comprise central government bonds, regional and local government bonds, as well as some social security funds.

[11] Kaminska and Zinna (2020) use the amount of Treasuries held by arbitrageurs as scaling variable.

The APP initially consisted of three components: the PSPP, an asset-backed securities purchase program (ABSPP), and a covered bond purchase program (CBPP3). A fourth component was added with the corporate sector purchase program (CSPP) in March 2016. The PSPP is by far the largest component, making up 84 percent of total net purchases, against 8 percent in the CBPP3, 7 percent in the CSPP, and 1 percent in the ABSPP. Within the PSPP, 90 percent of purchases (88 percent until March 2016) are made in national sovereign bonds, while 10 percent (12 percent) are allocated to euro area supranational issuers. The allocation of purchases across national bond markets is guided by the subscription of the 19 euro area national central banks (NCBs) in the ECB's capital key.[12]

By focusing on the general government debt of the "big-four" euro area countries for our measure of duration risk, we abstract from the remaining 15 euro area countries which account for the remaining 20 percent of euro area debt, the purchase of other agencies and supranational bonds, as well as private sector purchases within the ABSPP, the CSPP, and the CBPP3. This aligns the range of securities included in our supply variable with the GDP-weighted yields of the "big-four" euro area countries, which is our variable of interest. While theory suggests that purchases of corporate bonds would likewise decrease the overall duration risk to be borne by the market and hence affect the term premium of government bonds, in practice (due to market fragmentation and the lower liquidity in the non-big-four bond markets) one may expect that the cross-market impacts of such purchases are more muted than those happening in the same issuer universe. Taking such differentiated effects into account would require incorporating a notion of market fragmentation into our model, which goes beyond the scope of this paper. Conversely, including non-big-four sovereign bonds and corporate bonds in the yield measure to align it with a supply measure that includes

[12]The so-called ECB capital key refers to the subscription shares by the euro area NCBs in the capital of the ECB. The capital key subscription reflects the share of the respective member states in the total population and gross domestic product of the euro area, in equal measure.

duration risk from non-big-four sovereign bonds and corporate bonds would introduce non-negligible liquidity premia and credit risk premia into the yield measure. Moreover, importantly, our measure of the free-float of duration risk is a ratio, which captures the free-float of duration borne by arbitrageurs relative to total duration supply. Considering the private sector purchases and supply as well as the remaining 15 jurisdictions should leave the evolution of this ratio essentially unchanged, as both the numerator and the denominator of this ratio would adjust in that case. Finally, an advantage of our focus on the general government bonds of the big-four euro area countries is that for this set of securities it allows us to construct a granular and accurate measure of the free-float of duration risk. By contrast, using also the data for the remaining 15 countries, as well as private sector assets, would come at considerable computational cost and could introduce measurement errors, as the sectoral holdings data, which we would exploit to identify the holdings by arbitrageurs, require significant data cleaning.

2.2 Projecting APP Duration Extraction over Time

Central bank asset purchases exert their impact on the term structure by reducing the free-float of duration risk to be borne by arbitrageurs. Importantly, the theoretical model by Vayanos and Vila (2021) implies that the yield impact of central bank asset purchases in a specific maturity spectrum depends on the evolution of the discounted duration of the stock of bonds held by the central bank over the entire life of bonds in this spectrum. Therefore, beyond measuring the contemporaneous free-float of duration risk, we also need to project, at any given point in time, the free-float of duration risk, and its reduction through central bank asset purchases, into the future.

Projecting the evolution of the duration-weighted central bank portfolio requires information on future purchase volumes. We use the fact that the ECB's forward guidance on the path of net asset purchases was communicated in terms of an intended monthly purchase pace and horizon. For example, at the initial announcement of the APP in January 2015, the ECB Governing Council communicated its intention to make net purchases of €60 billion a month from

March 2015 to at least September 2016. After this initial announcement, the Governing Council made changes to the purchase horizon and/or the size of monthly flows in December 2015, March 2016, December 2016, October 2017, and June 2018. Each of these dates provides an "APP vintage," which is associated with a specific announced path for net purchases.[13]

In addition, we assume that announced net purchases are wound down along a linear tapering path, which reflects the ECB's early guidance that net purchases would not end abruptly. The linear tapering is assumed to reduce the monthly net purchase volume from the announced end-date in steps of €10 billion. Moreover, in December 2015 it was announced that maturing principals would be reinvested "for as long as necessary." From then on we assume a reinvestment phase to follow net asset purchases. From December 2015 to October 2017 we assume the reinvestment horizon to be two years. This is in line with median survey-based reinvestment expectation in the December 2017 Bloomberg survey, which first recorded reinvestment expectations. For June 2018, we use a median reinvestment horizon of three years as recorded in the respective Bloomberg survey. Table 2 summarizes, under the label "GovC," the key parameters for the various APP vintages.

Projecting the duration-weighted central bank portfolio also requires information on the maturity distribution of purchases. As announced in January 2015, securities with maturity of 2 to 30 years were eligible for purchase. Within this spectrum, the Governing Council communicated that purchases would be made in a "market-neutral manner." "Market neutrality" is understood to mean that the maturity distribution of the monthly flow of purchases is proportional to the eligible bond universe. Furthermore, initially, no purchases of securities with a yield below the deposit facility rate were undertaken. This constraint was relaxed in December 2016 from when purchases of securities with a yield below the deposit facility rate were allowed "to the extent necessary." At the same point, the

[13]The "time leg" of the forward guidance on asset purchases was complemented by a state-contingent forward-guidance element, according to which purchases would "in any case be conducted until the Governing Council would see a sustained adjustment in the path of inflation which is consistent with its aim of achieving inflation rates below, but close to, 2% over the medium term."

Table 2. APP Parameters Based on Governing Council Announcements and Survey Expectations

Date	Type	Monthly Pace (€bn) and Horizon		Total Net Purchases (€bn)	WAM (Years)
Jan. 22, 2015	GovC	60 Linear Taper	Mar. 2015–Sep. 2016 Oct. 2016–Feb. 2017	1,290	8.29
Mar. 12, 2015	Survey	60 Linear Taper	Mar. 2015–Sep. 2016 Oct. 2016–Feb. 2017	1,290	8.43
Nov. 25, 2015	Survey	60 75 Linear Taper	Mar. 2015–Dec. 2015 Jan. 2016–Mar. 2017 Apr. 2017–Oct. 2017	1,988	8.30
Dec. 3, 2015	GovC	60 Linear Taper Reinvestment	Mar. 2015–Mar. 2017 Apr. 2017–Aug. 2017 Sept. 2017–Aug. 2019	1,650	8.23
Mar. 3, 2016	Survey	60 80 Linear Taper Reinvestment	Mar. 2015–Mar. 2016 Apr. 2016–Mar. 2017 Apr. 2017–Oct. 2017 Nov. 2017–Oct. 2019	2,020	8.69
Mar. 10, 2016	GovC	60 80 Linear Taper Reinvestment	Mar. 2015–Mar. 2016 Apr. 2016–Mar. 2017 Apr. 2017–Oct. 2017 Nov. 2017–Oct. 2019	2,020	8.70
Dec. 2, 2016	Survey	60 80 72 Linear Taper Reinvestment	Mar. 2015–Mar. 2016 Apr. 2016–Mar. 2017 Apr. 2017–Aug. 2017 Sep. 2017–Apr. 2018 May 2018–Apr. 2020	2,388	8.49

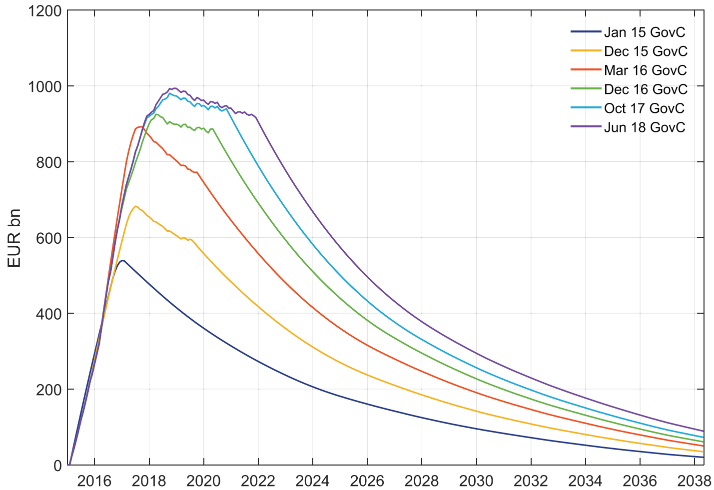
(continued)

Table 2. (Continued)

Date	Type	Monthly Pace (€bn) and Horizon		Total Net Purchases (€bn)	WAM (Years)
Dec. 8, 2016	GovC	60 80 60 Linear Taper Reinvestment	Mar. 2015–Mar. 2016 Apr. 2016–Mar. 2017 Apr. 2017–Dec. 2017 Jan. 2018–May 2018 Jun. 2018–May 2020	2,430	7.27
Oct. 18, 2017	Survey	60 80 60 Linear Taper Reinvestment	Mar. 2015–Mar. 2016 Apr. 2016–Mar. 2017 Apr. 2017–Dec. 2017 Jan. 2018–Sep. 2018 Oct. 2018–Sep. 2020	2,550	7.52
Oct. 26, 2017	GovC	60 80 60 30 Linear Taper Reinvestment	Mar. 2015–Mar. 2016 Apr. 2016–Mar. 2017 Apr. 2017–Dec. 2017 Jan. 2018–Sep. 2018 Oct. 2018–Nov. 2018 Dec. 2018–Nov. 2020	2,580	7.21
Jun. 7, 2018	Survey	60 80 60 30 Linear Taper Reinvestment	Mar. 2015–Mar. 2016 Apr. 2016–Mar. 2017 Apr. 2017–Dec. 2017 Jan. 2018–Sep. 2018 Oct. 2018–Dec. 2018 Jan. 2019–Dec. 2021	2,593	7.24
Jun. 14, 2018	GovC	60 80 60 30 15 Reinvestment	Mar. 2015–Mar. 2016 Apr. 2016–Mar. 2017 Apr. 2017–Dec. 2017 Jan. 2018–Sep. 2018 Oct. 2018–Dec. 2018 Jan. 2019–Dec. 2021	2,595	7.24

Note: The table reports the main APP purchase parameters for the different Governing Council announcements ("GovC"), as well as based on market expectations as reflected in Bloomberg surveys ("Survey"). The WAM reported is the WAM at the end of the net purchases associated. For each GovC series, a linear tapering is added to the announced pace and horizon, in line with the GovC communication that net purchases would not end abruptly. The survey parameters reported reflect the median responses. For Mar. 2015, Nov. 2015, and Mar. 2016 only the expected end date of net purchases is reported in the survey, and the linear tapering is added by assumption. Following the announcement at the Dec. 2015 GovC that maturing principals would be reinvested "for as long as necessary," a reinvestment phase is assumed. From Dec. 2015 to Oct. 2017, the assumed reinvestment phase is two years, in line with median survey-based reinvestment expectation in the Dec. 2017 survey, which first recorded such expectations. In Jun. 2018, the median reinvestment expectation was three years.

Figure 3. Evolution of ECB's Duration-Weighted Government Bond Holdings



Note: The figure shows the projected evolution of the government bond holdings for the big-four euro area countries in terms of 10-year equivalents. The different paths correspond to the APP vintages summarized in Table 2.

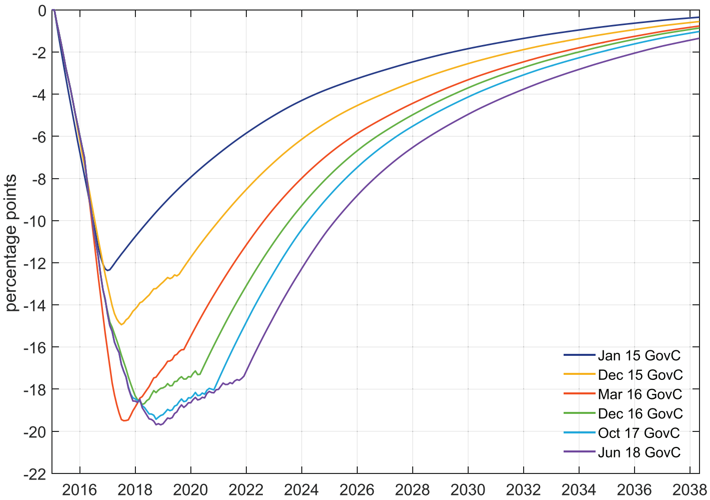
eligible maturity spectrum was extended from an interval spanning 2 to 30 years to a wider range of 1 to 30 years.

In addition, the projections account for further eligibility and operational criteria, which guided the implementation of historical purchases and affect their composition along several dimensions. First, the distribution of purchases across countries is determined by the ECB's capital key. Second, purchases respect an issue and issuer limit of 33 percent. Finally, future issuance of purchaseable securities is taken into account and based on the debt projections that enter the ECB's quarterly staff macroeconomic projection exercises and that were available at the given point in time.[14] Figure 3 shows the resulting projections of the APP portfolio in terms of 10-year equivalents for the different APP vintages summarized in Table 2.

Finally, we construct the trajectory of the free-float of duration risk over time by complementing the projections of the Eurosystem

[14]For details on the aggregate debt projections, see Bouabdallah et al. (2017).

Figure 4. Reduction in the Free-Float for Different APP Vintages



Note: The figure shows the compression of the free-float measure induced over time by the successive vintages of the APP. Each line shows the reduction of the free-float relative to the counterfactual of no APP. The different vintages are summarized in Table 2.

duration absorption with projections for duration supply. The projected duration supply is again based on the debt projections that enter the Eurosystem staff projections at a given point and are hence revised over time. In addition, we make the assumption that the WAM of the market portfolio remains unchanged over the projection horizon at the last observed WAM in each bond market. Figure 4 illustrates the compression of the free-float measure induced over time by the different vintages of the APP, constructed under the assumption that the ECB purchases reduce exclusively the holdings of price-sensitive investors. Each line shows the reduction of the free-float relative to the counterfactual of no APP.

Below we also study the effect of announcements on asset purchases on the yield curve. These should have an impact on sovereign bond yields only to the extent that they are unanticipated by financial markets. To quantify the yield impact of the initial APP announcement, as well as subsequent recalibration vintages, we therefore isolate the surprise in terms of additional future duration absorption associated with each announcement relative to what is already priced in based on pre-announcement market expectations, similar to Ihrig et al. (2018). We exploit the regular surveys by Bloomberg to obtain market expectations on the future purchase path of the APP. The resulting parameters are summarized in Table 2 under the label "Survey." We show for every APP recalibration the corresponding market expectations ahead of the recalibration announcement. In addition, we report the March 2015 survey path, as we use this in the estimation of our model (see Section 4.1).

The Bloomberg surveys were conducted systematically every six weeks from March 2015, and are typically published in the days ahead of the ECB Governing Council meetings.[15] The March 2015, November 2015, and December 2016 surveys did not contain information on expected "tapering" volumes. In those cases, we assume a linear tapering. The December 2016 survey contained information on expected tapering volumes, which we take into account. The October 2017 and June 2018 surveys provided a fully specified path for net asset purchases. Starting from the Governing Council's December 2015 reinvestment announcement until October 2017, we use a two-year reinvestment phase for the survey-based APP projections, in line with the Bloomberg survey of December 2017. For June 2018, we use a three-year reinvestment horizon in line with the corresponding survey. The maturity distribution of purchases for the survey vintages is assumed to follow the market-neutrality principle, in line with the approach taken for the Governing Council vintages. Using the survey-based information on the expected path of the APP, as well as assuming a market neutral maturity distribution of purchases, we create projections of the evolution of the market-expected duration-weighted APP portfolio, and the implied reduction of the free-float of duration risk.

2.3 Yields

In contrast to the U.S. Treasury market, there is no single sovereign fixed-income market at the level of the euro area as a whole.

[15]The March 2015 Governing Council meeting and survey provide an exception: the March 12, 2015 survey was conducted and published after the March 5, 2015 Governing Council meeting.

Each individual sovereign issues its own bonds. In order to provide a good representation of the overall sovereign debt market of euro area countries, we focus on the dynamics of the synthetic big-four euro area sovereign yield curve, which we construct as the GDP-weighted average of zero-coupon yields of Germany, France, Italy, and Spain.[16] The country-specific zero-coupon yields are constructed from prices of nominal bonds reported on the MTS platform. Bond prices are converted to zero-coupon yields using the Nelson-Siegel-Svensson methodology.[17] Our econometric analysis starts in December 2009 (in line with the availability of our SHS data) and ends in June 2018, when the Governing Council first expressed its anticipation to cease asset purchases by the end of December 2018, which was then subsequently confirmed. Figure 5 shows daily time series of the synthetic big-four zero-coupon yields for selected maturities over this period.

3. The Model

3.1 A Term Structure Model with Quantities

For tracing the impact of the APP on the yield curve, we rely on the model introduced by Li and Wei (2013). Yield curve dynamics are parsimoniously captured by three observable factors. The first two factors are given by the first two principal components (PCs) extracted from a cross-section of observed yields; see details in Section 4.1. We denote the first PC as the level factor L_t and the second PC as the slope factor S_t. The third factor, Q_t, is our free-float measure; see Equation (1). We collect the three factors in the vector $X_t = (L_t, S_t, Q_t)'$.

The short-term interest rate i_t is a linear combination of the factors

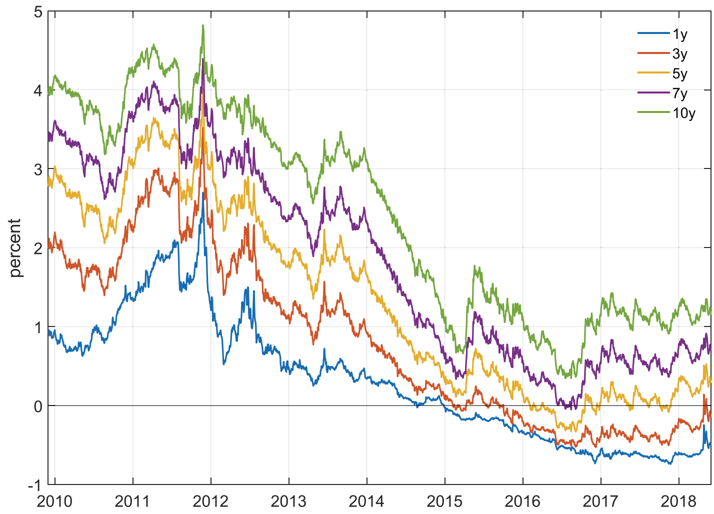
$$i_t = \delta_0 + \delta_1' X_t, \tag{2}$$

where we impose the constraint that $\delta_1 = (\delta_{1L}, \delta_{1S}, 0)'$, i.e., as in Li and Wei (2013), Q_t does not affect the short rate contemporaneously.

[16] The weights for Germany, France, Italy, Spain, are 0.38, 0.27, 0.21, 0.14, respectively.

[17] Information about the ECB's methodology for deriving zero-coupon yields is available at https://www.ecb.europa.eu/stats/financial_markets_and_interest_rates/euro_area_yield_curves/html/index.en.html.

Figure 5. Euro Area Zero-Coupon Yields



Note: The figure displays daily time series of the GDP-weighted synthetic zero-coupon yields for selected maturities of the big-four euro area countries for selected maturities from December 2009 to June 2018.

The factors X_t follow a VAR(1),

$$X_t = c + \mathcal{K}X_{t-1} + \Omega\epsilon_t, \quad \epsilon_t \sim N(0, I). \tag{3}$$

Following Li and Wei (2013), we constrain the autoregressive matrix $\mathcal{K}$ to be block-diagonal with the two blocks (L_t, S_t) and Q_t. In addition, the contemporaneous shock impact matrix Ω is assumed to be lower triangular:

$$\mathcal{K} = \begin{pmatrix} \mathcal{K}_{LL} & \mathcal{K}_{LS} & 0 \\ \mathcal{K}_{SL} & \mathcal{K}_{SS} & 0 \\ 0 & 0 & \mathcal{K}_{QQ} \end{pmatrix}, \quad \Omega = \begin{pmatrix} \Omega_{LL} & 0 & 0 \\ \Omega_{SL} & \Omega_{SS} & 0 \\ \Omega_{QL} & \Omega_{QS} & \Omega_{QQ} \end{pmatrix}. \tag{4}$$

Together with the assumption that the last element of δ_1 is zero, Equation (4) implies that the free-float measure Q_t does not forecast the short rate. In other words, the model excludes a potential signaling channel of central bank asset purchases.[18] In addition, the quantity measure is assumed not to be predictable by the yield curve factors.

The pricing kernel M_t is exponentially affine in the factors

$$M_{t+1} = \exp\left(-i_t - 0.5\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}\right), \tag{5}$$

where the market prices of risk λ_t are also affine in the factors

$$\lambda_t = \lambda_0 + \Lambda_1 X_t. \tag{6}$$

As in Li and Wei (2013) we impose the following zero constraints on the risk-compensation parameters λ_0 and Λ_1:

$$\lambda_0 = \begin{pmatrix} \lambda_{0,L} \\ \lambda_{0,S} \\ 0 \end{pmatrix}, \quad \Lambda_1 = \begin{pmatrix} \Lambda_{1,LL} & \Lambda_{1,LS} & \Lambda_{1,LQ} \\ \Lambda_{1,SL} & \Lambda_{1,SS} & \Lambda_{1,SQ} \\ 0 & 0 & 0 \end{pmatrix}. \tag{7}$$

This means that we assume that only level and slope risk is priced, but that the corresponding risk prices are driven by all three factors, including the quantity variable. Innovations to the quantity variable themselves are not priced, i.e., their market price of risk is zero.[19]

The market price of risk vector λ_t is where the effects of central bank asset purchases are determined in the model: changes in the quantity variable affect risk prices of level and slope risk and thereby term premia. The economic interpretation of this link between bond supply and term premia is discussed in more detail in Section 3.3 below.

Zero-coupon bond prices P_t^n of bonds with maturity n satisfy the no-arbitrage pricing equation

$$P_t^n = E\left(M_{t+1}P_{t+1}^{n-1}|X_t\right), \quad P_t^0 = 1. \tag{8}$$

[18]We worked with specifications that relax the zero restrictions in the short-rate equation (2) and the autoregressive matrix in (3) in order to allow for an impact of the free-float on (physical) short-rate expectations and hence for a signaling channel of QE to be active. However, those specifications did not lead to meaningful results (they were either insignificant or wrong-signed), so we did not pursue this route further.

[19]Li and Wei (2013) argue that "Treasury supply is unlikely to be a source of undiversifiable risk that should be priced on its own." Moreover, the zero restrictions on λ_0 and Λ_1 help to establish the analytical solutions for bond prices under given ("perfect foresight") investor expectations of future bond holdings as introduced in the following section.

Bond prices are converted into yields via $y_t^n = -\frac{1}{n} \log P_t^n$. Given the affine structure of the model, yields turn out to be affine functions of X_t,

$$y_t^n = -\frac{1}{n} \left(A_n + B_n' X_t \right), \tag{9}$$

where A_n and B_n satisfy the usual difference equations

$$A_{n+1} = A_n + B_n'(c - \Omega\lambda_0) + \frac{1}{2} B_n' \Omega \Omega' B_n - \delta_0, \tag{10}$$

$$B_{n+1}' = B_n'(\mathcal{K} - \Omega\Lambda_1) - \delta_1', \tag{11}$$

with $A_0 = 0$ and $B_0 = 0$. If $\Lambda_{1,LQ} \neq 0$ and $\Lambda_{1,SQ} \neq 0$, then $B_{n,Q}$, the third element of B_n, is different from zero. Hence, the quantity variable affects bond yields through the impact on the risk-compensation parameters but does not have a contemporaneous effect on the short rate.[20]

3.2 Bond Supply and the Yield Curve: Modeling Anticipated Shocks

How does the model translate a change in central bank asset purchases into changes in term premia and bond yields?

Under the standard approach in affine term structure models, time-t bond yields can only change if the time-t factors change. Iterating the no-arbitrage bond pricing Equation (8) forward, one obtains

$$P_t^n = E \left(M_{t+1} \cdot M_{t+2} \cdot \ldots \cdot M_{t+n} | X_t \right). \tag{12}$$

Bond prices (and thus yields) depend on the expected sequence of future pricing kernels (short rates and risk compensation), which

[20] The quantity variable does forecast level and slope factors—and hence future short-term rates via (2)—under the so-called risk-neutral dynamics. In a nutshell, under the risk-neutral probability measure, denoted by $\mathbb{Q}$, the pricing equation (8) boils down to discounting with the short-term rate (as opposed to the pricing kernel (5) involving risk corrections), i.e., $P_t^n = E^{\mathbb{Q}}(e^{-i_t} P_{t+1}^{n-1} | X_t)$, and incorporating the risk correction implicitly by using an amended factor process, which under the $\mathbb{Q}$ measure follows $X_t = \tilde{c} + \tilde{\mathcal{K}} X_{t-1} + \Omega \epsilon_t^Q$, where $\tilde{c} = c - \Omega\lambda_0$ and $\tilde{\mathcal{K}} = \mathcal{K} - \Omega\Lambda_1$.

are, in turn, a function of future state variables and their innovations. Thus, current yields depend on the full path of our free-float measure over the lifetime of the bond. At the same time, the information set in Equation (12) is just the current state variables X_t which are driven by a VAR. As a result, expectations of future state variables (including Q) and pricing kernels can only change if the current state variables X_t change. Accordingly, as indicated by the closed-form solution of the model (9), bond yields change if and only if current factors change. In particular, changes to central bank asset purchases could only be captured by a change to current Q_t. This change would trigger a change of future expected Q_{t+h} and future risk pricing—via Equation (6)—but this change in expectations of future quantities is fully determined by the change in the current (time-t) quantity.

However, the rigid link between future expectations and current state variables—implied by the standard approach described before—does not square well with the empirical evidence. In practice, the pure announcement of central bank asset purchases (i.e., statements affecting future Q_{t+h}) can have a significant impact on the yield curve today without contemporaneously moving Q_t at all. Moreover, while asset purchases are ongoing, too, further announced and credible changes to future purchase parameters—for example, a prolongation of the reinvestment horizon—can affect current bond yields even if they do not contemporaneously affect Q_t. Finally, even if an innovation to the path of central bank asset purchases does affect the current free-float Q_t, the announced future changes to Q may differ from those implied by the conditional expectations $E_t(Q_{t+h})$ as prescribed by the VAR in Equation (3).

To capture the possibility that anticipated future free-float changes have an impact on the current yield curve over and beyond what is implied by the current states, Li and Wei (2013) allow anticipated innovations to the quantity variable to enter the bond pricing equation. Specifically, their approach amounts to conditioning bond pricing not only on current state variables, as in Equation (8), but also on a sequence of anticipated future free-float ratios $\bar{\mathcal{Q}}_t = \{\bar{Q}_t, \bar{Q}_{t+1}, \bar{Q}_{t+2}, \dots\}$

$$P_t^n = E\left(M_{t+1} P_{t+1}^{n-1} | X_t, \bar{\mathcal{Q}}_t\right), \quad P_t^0 = 1. \tag{13}$$

This is the same pricing expression as in Equation (8), except that it uses an enhanced set of conditioning information. Denote by $\mathcal{Q}_t^0 = \{Q_t^0, Q_{t+1}^0, Q_{t+2}^0, \ldots\}$ the sequence of expected free-float ratios based on the state vector X_t and the VAR dynamics in Equation (3). Let $u_{t+h} = \bar{Q}_{t+h} - Q_{t+h}^0$ denote the anticipated innovation of the free-float to the "baseline" and $\mathcal{U}_t = \{u_t, u_{t+1}, u_{t+2}, \ldots\}$ the corresponding sequence of anticipated innovations. Li and Wei (2013) show that bonds priced under the enhanced information set in Equation (13) satisfy the yield equation:

$$y_t^n = -\frac{1}{n} A_n + dy_n(\mathcal{U}_t) - \frac{1}{n} B_n' X_t, \tag{14}$$

where A_n and B_n are the same expressions as in Equation (9) in the standard setup and

$$dy_n(\mathcal{U}_t) = -\frac{1}{n} \left[B_{n,Q} u_t + \sum_{h=1}^{n} B_{n-h,Q} (u_{t+h} - \mathcal{K}_{QQ} u_{t+h-1}) \right]. \tag{15}$$

The expression $dy_n(\mathcal{U}_t)$ is the impact of a sequence of anticipated innovations to the quantity factor on the n-period term premium and corresponding yield over and beyond what is incorporated in current factors X_t.[21]

We can alternatively rearrange the terms in Equation (15) to obtain

$$dy_n(\mathcal{U}_t) = \sum_{h=1}^{n} \gamma_h^n u_{t+h-1} = \gamma^{n\prime} \mathcal{U}_t, \tag{16}$$

where $\gamma^n = (\gamma_1^n, \ldots, \gamma_n^n)'$ with

$$\gamma_h^n = -\frac{1}{n} \left(B_{n-h+1,Q} - \mathcal{K}_{QQ} B_{n-h,Q} \right). \tag{17}$$

[21] We consistently use $\mathcal{U}$ sequences that are longer than the lifetime of any bond. Therefore, we do not need to obey the distinction in Li and Wei (2013) regarding the upper summation limit that becomes relevant if $\mathcal{U}$ sequences are shorter than bond maturities. In terms of notation, in any scalar product involving $\mathcal{U}$, such as, e.g., in (16), $\mathcal{U}$ is assumed to have the same length as the corresponding multiplying vector.

The model stipulates a linear relationship between changes in the trajectory of the anticipated future free-float over the tenor of a bond and the change in the yield of that bond. The sensitivity of yields to anticipated innovations in the future free-float are captured by maturity- and horizon-specific "impact factors" γ_h^n, which are a function of the model parameters, such as the persistence of factors, innovation volatility, and market prices of risk.

In Section 5 we deploy Equation (16) to investigate how APP recalibrations have affected the yield curve. A certain APP surprise at time t is summarized by a corresponding $\mathcal{U}_t$ sequence and the yield impact is obtained via (16).

3.3 Bond Supply and the Yield Curve: Interpreting the Transmission Channel

In this section we provide further detail on the transmission channel of shocks to the free-float of duration risk implied by central bank asset purchases in our empirical model. First, we recall the standard decomposition of yields and show how expected future free-float measures $E(Q_{t+h}|X_t)$—with expectations being fully determined by current states—affect expected future excess returns and hence term premia. Second, we show that the same transmission channel holds for anticipated shocks, i.e., free-float innovations u_{t+h} that are *not* implied by contemporaneous state variables: we find that the effect of such an anticipated free-float shock u_{t+h} on future expected excess returns and hence term premia is the same as the effect stemming from a change in expected free-float induced by a change in current states, i.e., $E(Q_{t+h}|X_t)$.

The n-period bond yield can be represented as the sum of the expectations component (average expected future short rates over the lifetime of the bond) and the term premium. The term premium component is, in turn, given by the average of expected future excess returns:

$$y_t^n = \underbrace{\frac{1}{n} E_t \sum_{h=0}^{n-1} i_{t+h}}_{\text{Expectations component}} + \underbrace{\frac{1}{n} E_t \sum_{h=1}^{n} rx_{t+h}^{n-h}}_{\text{Term premium}}, \qquad (18)$$

where $rx^{n-h}_{t+h} = \ln P^{n-h}_{t+h} - \ln P^{n-h+1}_{t+h-1} - i_{t+h-1}$ is the one-period excess return for a bond with maturity $n-h+1$ purchased at time $t+h-1$. Unless specified otherwise, the conditional expectation $E_\tau(\cdot)$ is equivalent to $E(\cdot|X_\tau)$.

The identity in (18) is independent of a specific model. Different term structure models imply different parametric expressions for the expectations component and the term premium. For the affine model introduced in Section 3.1, each expected future excess return, conditional on information at the time of the purchase of the bond, can be expressed as[22]

$$E_{t+h-1} rx^{n-h}_{t+h} = B'_{n-h}\Omega\lambda_{t+h-1} + JI. \tag{19}$$

The term $B'_{n-h}\Omega$ captures factor sensitivity or "duration risk," i.e., the exposure of (log) bond prices to unexpected changes in risk factors, while λ_{t+h-1} is the time-varying "price of risk," i.e., the amount of excess return compensation per unit of risk. This compensation varies over time but is the same across bonds of all maturities, thus excluding arbitrage opportunities. The last item is a convexity adjustment (Jensen inequality) term, given by $JI = -0.5B'_{n-h}\Omega\Omega' B_{n-h}$, which does not depend on the factors.

The zero restrictions in (7) imply that level and slope risk is priced, while the risk at any future time of unexpected changes in the free-float measure is not priced, i.e., $\lambda_{t+h-1,Q} \equiv 0$. Grouping the level and slope factor as $Z_t = (L_t, S_t)'$ and writing the corresponding model matrices in a partitioned fashion (denoting the upper 2×2 part of Λ_1 in Equation (7) by $\Lambda_{1,ZZ}$, etc.), the market price of level/slope risk, i.e., the 2×1 vector $\lambda_{t+h-1,Z}$, is given by

$$\lambda_{t+h-1,Z} = \lambda_{0,Z} + \Lambda_{1,ZZ}Z_{t+h-1} + \Lambda_{1,ZQ}Q_{t+h-1}. \tag{20}$$

The time variation in the market price of level/slope risk is driven by the level and slope itself ($\Lambda_{1,ZZ}Z_{t+h-1}$) as well as by the free-float measure ($\Lambda_{1,ZQ}Q_{t+h-1}$). Rewriting (19) we obtain

[22] The link between Equations (18) and (19) can be seen by conditioning, in (19), the future expected one-period excess return on information (factors) at time t and applying the law of iterated expectations.

$$E_{t+h-1} rx_{t+h}^{n-h} = \underbrace{\Theta_{n-h} + \theta'_{n-h} Z_{t+h-1}}_{\text{Terms independent of } Q} + \underbrace{(B'_{n-h,Z}\Omega_{ZZ} + B_{n-h,Q}\Omega_{QZ})}_{\text{Exposure to level/slope risk}} \underbrace{\Lambda_{1,ZQ} Q_{t+h-1}}_{\text{Price of risk components driven by } Q}, \tag{21}$$

where Θ_{n-h} is a constant comprising the Jensen term JI in (19) and the time-invariant risk compensation (as function of $\lambda_{0,Z}$). The term $\theta'_{n-h} Z_{t+h-1}$ depends on the level and slope factors but not on the free-float measure. The last summand is the time-varying contribution of the quantity factor to the one-period expected excess return h–1 periods ahead. The first part in parentheses ("Exposure to level/slope risk") is the log bond price sensitivity to $\epsilon_Z = (\epsilon_L, \epsilon_S)'$ shocks in (3). This part affects the level and slope factors, and hence bond prices, either directly, via $B'_{n-h,Z}\Omega_{ZZ}$, or indirectly, by contemporaneously affecting the Q factor (via Ω_{QZ}) and affecting bond prices via the respective factor loading $B_{n-h,Q}$. The second part ("Price of risk components driven by Q") is the time-varying contribution of Q_{t+h-1} to the respective prices of level and slope risk. Conditioning Equation (19) on information at time t, we note that if the current free-float Q_t changes, this affects $E_t(Q_{t+h-1})$ via the VAR, which in turn shifts expected future excess returns at time t through a change in the expected market price of risk and thus the time-t term premium.

Having shown how expected free-floats—with expectations spanned by current state variables—have an impact on term premia in the standard approach of affine term structure models, we now explain that the same transmission channel holds for anticipated shocks, i.e., free-float innovations u_{t+h} that are not implied by contemporaneous state variables.

Recall that the impact factors γ_h^n in Equation (16) are expressed in terms of factor loadings on the free-float factor $B_{.,Q}$; see Equation (17). We now convert them into an alternative expression that highlights their economic interpretation as risk premium contribution. Starting from the recursion in Equation (11) and defining a selection vector $s = (0, 0, 1)'$, the impact of Q on the m-maturity log bond price is given by

$$B_{m,Q} = B'_{m-1}\mathcal{K}s - B'_{m-1}\Omega\Lambda_1 s - \delta'_1 s. \tag{22}$$

Grouping again $Z_t = (L_t, S_t)'$, partitioning system matrices accordingly, and noting the zero restrictions in $\mathcal{K}$, Ω, and Λ_1, we obtain

$$\mathcal{K} \cdot s = \mathcal{K}_{QQ} \cdot s, \quad \Omega\Lambda_1 s = \begin{pmatrix} \Omega_{ZZ}\Lambda_{1,ZQ} \\ \Omega_{QZ}\Lambda_{1,ZQ} \end{pmatrix}, \quad \delta_1' s = 0.$$

Therefore,

$$B_{m,Q} = B_{m-1,Q}\mathcal{K}_{QQ} - \left(B'_{m-1,Z}\Omega_{ZZ} + B_{m-1,Q}\Omega_{QZ}\right)\Lambda_{1,ZQ}.$$

Rewriting this expression for $m = n - h + 1$, we obtain from Equation (16) the expression for the impact factors

$$\gamma_h^n = -\frac{1}{n}\left(B'_{n-h,Z}\Omega_{ZZ} + B_{n-h,Q}\Omega_{QZ}\right)\Lambda_{1,ZQ}. \tag{23}$$

This is the same expression as in the last line of Equation (21). Therefore, an anticipated innovation u_{t+h-1} to the free-float has the same impact on the term premium as a change in the expected future free-float $E(Q_{t+h-1}|X_t)$ due to a change in current Q_t.

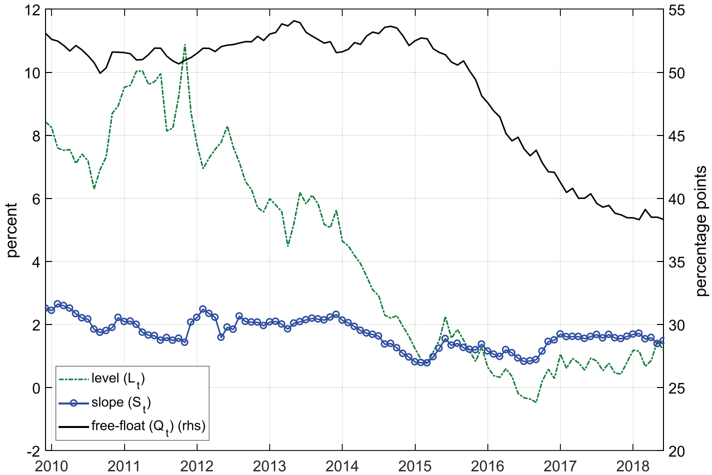
Overall, the model used in this paper is inheriting key features from the equilibrium model introduced by Greenwood and Vayanos (2014) and Vayanos and Vila (2021). A more detailed comparison between the empirical model and certain properties of their theoretical framework is provided in the appendix.

4. Estimation

4.1 Estimation Approach

While we rely on the same modeling framework as Li and Wei (2013), we modify their two-step estimation approach in order to address specific challenges posed by the euro area data. In the first step we estimate a VAR of the risk factors, including the free-float variable, and the relation between the short-term rate and these factors. In the second step, we quantify the market prices of risk by using a dual objective: we simultaneously match the time-series evolution of bond yields between December 2009 and August 2014, as well as the portion of the yield curve decline between September 2014 and March 2015 that can be attributed to markets gradually pricing in expectations for large-scale asset purchases by the Eurosystem.

Figure 6. Model Factors



Note: The figure shows end-of-month time series of the level and the slope factors (left axis) and the free-float factor expressed as a percentage of total duration supply (right axis).

Specifically, in the first step we fit a VAR(1) to an empirical level and slope factor (L_t and S_t, respectively) and to our observed free-float measure Q_t (see Equation (1)), over the pre-APP subperiod from December 2009 to August 2014. The level and the slope factors are extracted as the first two principal components from the cross-section of observed yields with maturities 1-year, 2-year, . . ., 10-year. Figure 6 shows monthly time series of the three factors over the full sample from December 2009 to June 2018. The shock impact matrix Ω is the Cholesky decomposition of the variance-covariance matrix of the reduced-form shocks implied by the estimated VAR model. We estimate the parameters δ_0 and δ_1 in Equation (2) with OLS. For the VAR and the OLS regression we impose the zero restrictions on $\mathcal{K}$ from Equation (4) and δ_1 from Equation (2), respectively.

In the second step, we estimate the market-price-of-risk parameters. In theory, we could follow Li and Wei (2013) and match the observed time series of bond yields and term premia estimates obtained from an auxiliary term structure model (Kim and Wright 2005) that excludes bond supply information. However, in practice two aspects of the euro area data prevent us from relying on such a pure time-series approach. First, our sample is relatively short due to the limited availability of the euro area free-float measure, which is available only from December 2009. Second, Eurosystem bond holdings only became a sizable source of variation in the free-float with the start of APP. By contrast, the Federal Reserve's SOMA (System Open Market Account) portfolio exhibited significant variations already before the inception of the Federal Reserve's LSAPs. Hence, based on the euro area data, it is more challenging for the model to learn about the parameters from the covariation of Q and bond yield dynamics.

Therefore, for estimating the market-price-of-risk parameters, our second step uses a dual objective function that not only takes into account the time-series fit of bond yields but also the model's ability to capture the initial decline of the yield curve from September 2014 to March 2015. From September 2014, expectations about a possible ECB future large-scale asset purchase program were building in financial markets ahead of the start of the APP in March 2015.

For the first part of the objective function, denote by y_t^o the cross-section of observed yields with maturities 1-year, 2-year, ..., 10-year, and by $\hat{y}_t \equiv \hat{y}_t(\lambda_0, \Lambda_1 | X_t; \hat{c}, \hat{\mathcal{K}}, \hat{\Omega}, \hat{\delta}_0, \hat{\delta}_1)$ the corresponding fitted yields, using Equation (9) and taking the estimated VAR and short-rate parameters from the first step as given. Our distance measure is the average (across maturities and time) squared fitting error using end-of-month yields from December 2009 to August 2014:

$$F_1(\lambda_0, \Lambda_1) = \frac{1}{M_1 T} \sum_{t=1}^{T} [y_t^o - \hat{y}_t]' [y_t^o - \hat{y}_t], \tag{24}$$

where $T = 56$ denotes the number of time-series observations and $M_1 = 10$ the number of maturities used in the cross-section.

For the second part of the objective function, we assume that part of the euro area bond yield decline observed from September 2014 to March 2015 (when the APP was officially launched) was due to the buildup of private sector expectations of large-scale sovereign bonds purchases. To infer the cumulative yield decline over this time window that can be attributed to the anticipation of the APP, we conduct an event study. We then match the observed APP-induced cumulative change in bond yields with the model-implied change in term premia, conditional on a proxy for the prevailing APP expectations at the time.

For the selection of events with APP-related news, we follow Altavilla, Carboni, and Motto (2021) and focus on a set of event dates at which the ECB conveyed news about the APP in the form of ECB press conferences as well as speeches given by ECB President Draghi. The first date is September 4, 2014, the day of the ECB press conference at which the initial purchases under the ABSPP and the CBPP3, which preceded the announcement of the APP in January 2015, were communicated. Moreover, at the same point, President Draghi indicated that a "broad asset purchase programme was discussed, and some Governors made clear that they would like to do more."[23] The last date is March 5, 2015, when the ECB announced final technical details of the program, which complemented the information provided at the press conference following the January 22, 2015 Governing Council, and which confirmed March 9, 2015 as the starting date for the APP.[24]

For our event study we analyze the changes of zero-coupon yields over two-day windows.[25] We assume that the observed changes in yields around those event dates are primarily driven by market participants' changing expectations about the APP. Following Altavilla, Carboni, and Motto (2021), we conduct two versions of the event study: one in which we control for news about key macroeconomic variables on those event dates, and another without such controls. We restrict our focus on the medium- and long-term segment of the yield curve and disregard changes of yields with less than five-year maturity. This is motivated by the fact that average short-rate expectations over shorter horizons may also reflect monetary policy news unrelated to the APP. From September 3, 2014

[23]See https://www.ecb.europa.eu/press/pressconf/2014/html/is140904.en.html.

[24]Our selection of relevant dates is alligned with the working version of Altavilla, Carboni, and Motto (2021) where March 9, 2015—the date when actual APP purchases started—is not included in the set of events.

[25]Most of the relevant ECB announcements were made in the afternoon on a given day. We consider two-day rather than one-day yield changes, as the construction of zero-coupon yields (see Section 2.3) for a given day may incorporate prices prevailing before noon. Thus, if the announcement took place at date t, some of the price changes underlying the construction of zero-coupon yields between $t-1$ and t may not reflect the event of day t's afternoon.

to March 6, 2015 zero-coupon bond yields declined by 89 bps at the 10-year maturity (and 46 bps at the 5-year maturity). Averaging the results of the two event-study analyses (controlled versus uncontrolled), we attribute cumulative reductions of the 10-year (and 5-year) zero-coupon bond yield of 48 bps (and 33 bps) to APP announcements.

To operationalize the second component of the objective function, let dy^o denote the change in bond yields for maturities 5-year, 6-year, . . ., 10-year, which is attributable to news about the APP as estimated by the aforementioned event-study approach. For example, $dy^o_{10y} = -48$ bps. Let $\hat{dy} \equiv \hat{dy}(\lambda_0, \Lambda_1 | \mathcal{U}_t; \hat{c}, \hat{\mathcal{K}}, \hat{\Omega}, \hat{\delta}_0, \hat{\delta_1})$ denote the corresponding model-implied changes over the same period, computed by deploying Equation (16) for the respective maturities. The $\mathcal{U}$ sequence used in (16) represents the expected trajectory for duration extraction determined by the APP as of March 5, 2015. This trajectory is constructed based on survey expectations prevailing at that date, which were closely aligned to the Governing Council's January 2015 announcements. We assume that this $\mathcal{U}$ sequence represents the APP expectations prevailing when the program was launched, as they had built up "from zero" from September 2014. The second part of the objective function is then given by the distance measure:

$$F_2(\lambda_0, \Lambda_1) = \frac{1}{M_2}[dy^o - \hat{dy}]'[dy^o - \hat{dy}], \tag{25}$$

where $M_2 = 6$ denotes the number of maturities used in the second part of the objective function.

The optimization problem for estimating the market-prices-of-risk parameters then is

$$\{\hat{\lambda}_0, \hat{\Lambda}_1\} = \underset{\{\lambda_0, \Lambda_1\}}{\arg\min}\ \omega F_1(\lambda_0, \Lambda_1) + (1 - \omega) F_2(\lambda_0, \Lambda_1), \tag{26}$$

where ω is a weighting parameter that balances the importance of the time-series fit criterion F_1 and the "event-window" fit criterion F_2 for the overall objective function. Choosing ω requires judgment. With a view of imposing a "flat prior" across the two criteria, we set $\omega = 0.5$. However, altering the weight ω does not affect our estimates of parameters and yield curve impacts very much, as it turns out that the events component mainly informs the market-price-of-risk parameters governing the mapping from free-float to term premia that are of relatively minor relevance for the overall time-series fit over the first part of our sample.

4.2 Parameter Estimates

Table 3 reports estimates of the model parameters. Using estimates of the price-of-risk parameters λ_0 and Λ_1, which are derived in the second estimation step, we compute estimates for the parameters $\tilde{c}$ and $\widetilde{\mathcal{K}}$ that govern the risk-neutral dynamics of factors X_t. The higher eigenvalues of $\widetilde{\mathcal{K}}$ than $\mathcal{K}$ indicate that all three factors are more persistent under the risk-neutral ($\mathbb{Q}$) than the real-world ($\mathbb{P}$) probability measure.

Given these parameter estimates and using the affine relation (9) between factors X_t and bond yields y_t^n, we report in Table 4 the reaction of the yield curve to a time-t increase in each factor, which amounts to one standard deviation of the reduced-form shocks derived from the estimate of the shock variance-covariance matrix $\Omega\Omega'$. The first column of Table 4 indicates that the loadings of yields on the level factor L_t are positive and of similar size across maturities. Thus, a positive shock to this risk factor leads to an (almost) parallel upward shift of the entire yield curve. A positive shock to the slope risk factor S_t leads to a steepening of the yield curve, as indicated by the second column of Table 4.

A contemporaneous shock to the quantity factor Q_t shifts the entire yield curve in the same direction as the shock, as indicated by the third column of Table 4. Therefore, in line with economic intuition, yields decrease when the free-float measure is reduced by the Eurosystem's duration extraction. The yield impact of a shock to the free-float is hump-shaped across maturities. This is one of two possible shapes that can arise in the equilibrium model by Greenwood and Vayanos (2014). As argued by the authors, the hump-shaped pattern can occur when the shock to the free-float is mean-reverting relatively quickly. Indeed, according to the estimated $\mathcal{K}_{QQ}$ of 0.9039 from Table 3—which represents the persistency of the Q factor due to the imposed restrictions on the interactions with the other two factors—the impact of a shock to the free-float has a half-life of only seven months. In addition, we find that a shock to the

Table 3. Parameter Estimates

	δ_0 0.0045			
	δ_1			
L_t S_t Q_t	0.1240 –0.4432 0			
		$\mathcal{K}$		c
L_t S_t Q_t eig ($\mathcal{K}$)	0.9835 0.0150 0 0.8256	–0.0748 0.8185 0 0.9039	0 0 0.9039 0.9764	0.0016 0.0024 0.0501
		Ω		
L_t S_t Q_t	0.0067 –0.0003 0.0001	0 0.0018 0.0002	0 0 0.0041	
		Λ_1		λ_0
L_t S_t Q_t	–5.1457 33.4122 0	–17.5141 –63.7912 0	–0.5450 1.4627 0	0.5768 –1.2922 0
		$\widetilde{\mathcal{K}}$		$\tilde{c}$
L_t S_t Q_t eig ($\widetilde{\mathcal{K}}$)	1.0181 –0.0450 –0.0062 0.9044	0.0429 0.9253 0.0132 0.9567	0.0037 –0.0027 0.9036 0.9860	–0.0023 0.0048 0.0504

Note: This table reports parameter estimates of the model obtained by using the two-step approach described in Section 4.1. In the first step we derive estimates for c, $\mathcal{K}$, and Ω—which govern the real-world dynamics of factors X_t—and of δ_0 and δ_1—which map linearly factors X_t into the short rate—and, in the second step, for market-price-of-risk parameters λ_0 and Λ_1. Given these estimates, we compute $\tilde{c} = c - \Omega\lambda_0$ and $\widetilde{\mathcal{K}} = \mathcal{K} - \Omega\Lambda_1$, which govern the risk-neutral dynamics of factors X_t.

free-float moves the contemporaneous one-period expected returns $E_t rx_{t+1}^{n-1} = B'_{n-1}\Omega\lambda_t$ of all bonds in the same direction as the shock, and that the effect is increasing across maturities. Also this empirical finding of our euro area model is in line with the prediction of the

Table 4. Reaction of the Yield Curve to Changes in the Factors

Maturity of Yield (Years)	L_t	S_t	Q_t
1	15	–5	0.29
2	20	–2	0.41
3	22	–1	0.44
4	23	0	0.44
5	23	1	0.42
6	23	1	0.41
7	23	2	0.38
8	22	2	0.36
9	21	2	0.34
10	20	2	0.32

Note: The table reports changes in yields in response to a one-standard-deviation shock in time-t to each factor in $X_t = (L_t, S_t, Q_t)'$. The changes are derived using Equation (9). The changes are reported in basis points.

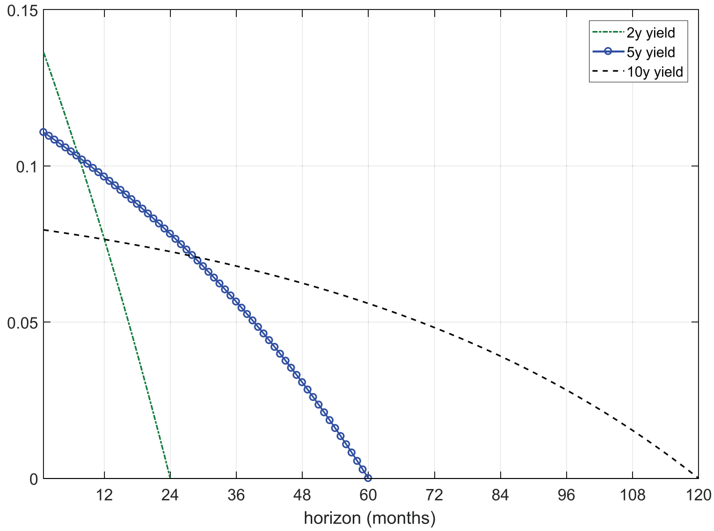
theoretical model of Greenwood and Vayanos (2014); see Section 1.3 of their paper.

Finally, we inspect the estimated impact factors γ that map sequences of innovations of the free-float from the expected path implied by the VAR dynamics—Equation (3)—into yield curve reactions. Figure 7 plots the estimated impact factors γ_h^n for maturities $n = 2$, 5, and 10 years over the relevant horizons h. The impact factors decrease monotonically over the future horizons within the tenor of any bond. Therefore, changes in free-floating duration supply over the near term have a larger effect on the term premium component of a yield than changes occurring in the more distant future. This pattern holds for bonds of any maturity.

4.3 Model Fit

For our time-series fitting criterion F_1 in (26), the model delivers a good fit to the yield-level data over time. Root mean squared errors over the estimation period range from 3 to 14 bps across maturities. This is comparable in size to the fit of U.S. yields in Li and Wei

Figure 7. Estimated Impact Factors

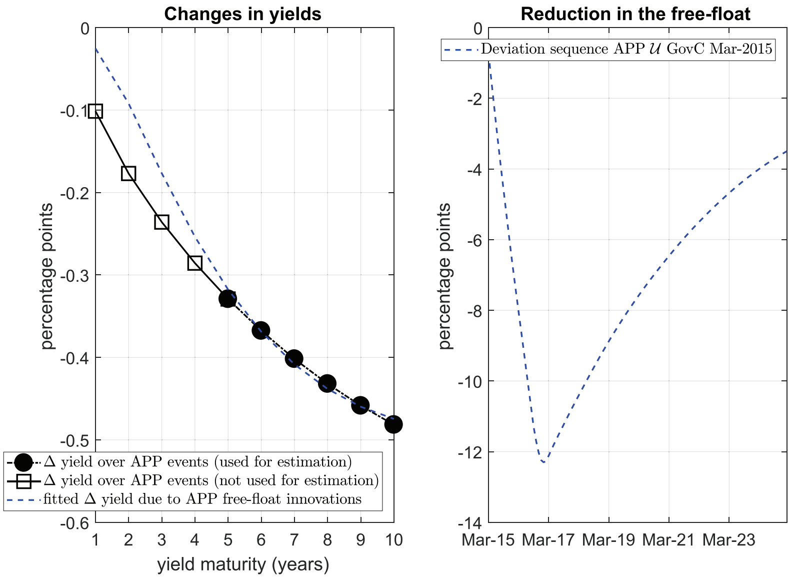


Note: The figure shows the estimated impact factors γ_h^n that map a sequence of revisions to the current and expected free-float into changes in yields of bonds with maturity $n = 2$, 5, and 10 years; see Equation (16). The vertical axis shows the yield change contribution per unit of free-float change at the respective horizon. For instance, an anticipated free-float reduction in three years by 1 percentage point contributes to lowering the five-year yield by 0.05 percentage point contemporaneously.

(2013). In particular, medium- and long-term yields are fitted well in sample.

Focusing on F_2, the second part of the objective function in (26), the left panel of Figure 8 presents the model fit of the APP-induced cumulative decrease in the sovereign yield curve over selected event dates from September 2014 to March 2015, as discussed in Section 4.1. The right panel of Figure 8 plots the series of anticipated future APP-induced free-float innovations $\mathcal{U}$ as of March 5, 2015 that underlies these fitted yield changes. Each model-implied yield decline shown is the result of multiplying the maturity-specific impact factors γ_h^n—see Equation (16)—with the part of this $\mathcal{U}$ sequence spanning the horizons of each yield tenor. For example, the 10-year yield APP-induced compression (of about 49 bps) is the product of the γ_h^{10y}, the dashed line in Figure 7, and the part of the $\mathcal{U}$ sequence

Figure 8. Impact of the Anticipation of the APP on the Yield Curve through Expected Future Free-Float Compression



Note: The left panel plots observed and fitted changes in yields over an event window; the black line represents the cumulative decreases in yields over the APP-related events from early September 2014 to early March 2015; the blue dashed line plots the model-implied changes in yields due to future APP-induced free-float innovations $\mathcal{U}$ as of March 5, 2015 (shown in the right panel). Decreases in yields with maturities as of 5 years up to 10 years (circle markers) are used in the estimation of the model, while the decreases of yields with shorter maturities (square markers) are left out.

from the right panel of Figure 8 starting in March 2015 and ending in February 2024.

The model fits almost perfectly the decreases in yields with maturities of five years and more, which corresponds to the data used in the second part of the objective function. For shorter maturities, which do not enter the estimation criterion, the model predicts less pronounced yield decreases than observed for the selected events. As the model captures only the effect of the APP on term premia due to the duration extraction channel, the observed undershooting of the model prediction is attributable to factors outside our model, such as a signaling channel of APP-related communication, or changes in the expected ECB's key interest rate policy rates.

The reduction in the future free-float induced by the anticipation of the APP—see the right panel of Figure 8—that underlies the fitted decreases in yields between September 2014 to March 2015 is large relative to the average variation of the supply factor from December 2009 to August 2014. The standard deviation of innovations to the free-float in this early period amounts to only 0.4 percentage point (see the estimate of the shock variance-covariance matrix $\Omega\Omega'$ from Table 3). By contrast, in March 2015 the anticipated reduction in the free-float induced by the APP was envisaged to peak at about 12 percentage points at the end of the net purchases phase and to still amount to about 4 percentage points in 2025; see again Figure 8. Hence, in contrast to the short-lived persistence of a shock to the free-float in the pre-APP period (with a half-life of only seven months; see Table 3), the APP represents a very persistent reduction in the supply of available bonds. In terms of implementation, recall that we condition in the bond pricing equations on arbitrary free-float processes—in turn based on ECB policy announcements and market participants' expectations gleaned from surveys—thereby "overwriting" the free-float dynamics implied by the estimated VAR. In the standard model setup without such feature, any free-float-induced change in yields would need to come via a change in the contemporaneous state variable, i.e., via an adequately sized innovation, and expectations of future free-float would follow from VAR dynamics. Here, following Li and Wei (2013), we can directly condition on a given free-float path; see again Equation (13). To illustrate further the extraordinary dimension of the APP compared with historical free-float variation, we can compute the unanticipated contemporaneous free-float shock that, when multiplied with the respective yield loading $B_{10y,Q}$, would give the same 49 bps impact on the 10-year yield as the anticipated free-float shock sequence as of March 2015. It turns out that this hypothetical shock would amount to 61 percentage points. This represents more than the actual supply of bonds available to arbitrageurs at the time the program was launched; see Figure 6. Overall, both the size and the persistence of APP-induced innovations to future free-float are several orders of magnitude higher than the free-float variation observed in the pre-APP sample.

5. The Impact of the APP on the Yield Curve

We use our estimated model to infer the impact of the APP on the sovereign yield curve through the duration extraction channel. First, we estimate the compression of term premia along the yield curve for different vintages of the APP (Section 5.1). Second, we examine the persistence of the term premium compression over time and investigate the contribution made by reinvestments of maturing principals (Section 5.2). Third, we compare the yield changes observed around APP recalibration announcements to the real-time predictions of our model (Section 5.3). Finally, we assess the robustness of our results (Section 5.4).

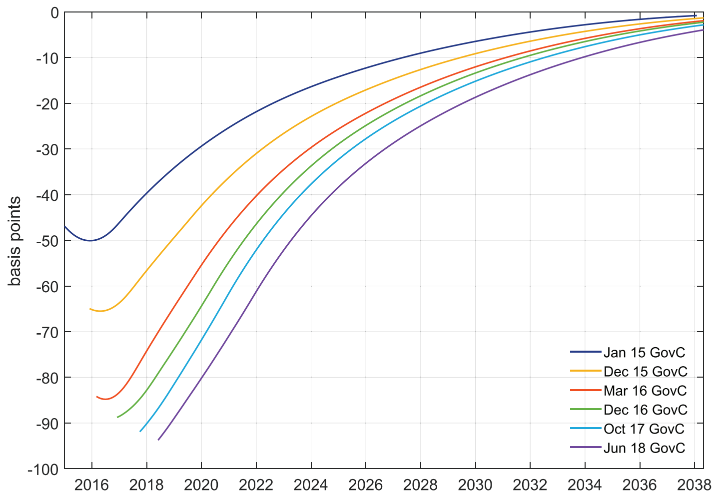
5.1 Term Premium Compression across the Yield Curve

Figure 1 shows the estimated impact of the APP across the yield curve for the different APP vintages at the time of their announcement. Each curve shows the estimated term premium compression relative to the counterfactual of no duration extraction through the APP. The date shown in the legend indicates both the respective APP vintage, i.e., the specific path of net purchases implied by the vintage, as well as the point in time at which the term premium compression is estimated. For example, the curve labeled "Jan 15 GovC" shows the estimated term premium compression due to the January 2015 APP vintage in January 2015.

We obtain these yield curve impact estimates by feeding the free-float reduction implied by the different APP vintages (see Figure 4) into our model. In detail, the estimated term premium compression is constructed using Equation (16), which maps the sequence of anticipated free-float innovations ($\mathcal{U}$) into a yield impact using the impact factors (see illustrations of γ_h^n for selected maturities in Figure 7). For the example of January 2015, the sequence of free-float innovations relevant for the 10-year term premium is the part of the dark-blue line corresponding to the January 2015 APP vintage in Figure 4 that ranges from January 2015 10 years to December 2024. To compute the five-year term premium, the relevant sequence of free-float innovations consists of just the five years from January 2015 until December 2019 in Figure 4.

For January 2015, the impact on the 5-year and 10-year term premium is found to be around 30 and 50 bps, respectively. Also for

Figure 9. The Impact of the APP on the 10-Year Term Premium over Time



Note: For selected dates, the figure shows the estimated contemporaneous and future impact of the APP on the term premium component of the 10-year sovereign bond yield.

the subsequent vintages, the term premium impact is estimated to be higher for longer tenors, i.e., the APP has led to a flattening of the curve. The overall term structure impact has become stronger over time as the APP has been expanded in length and volume. For the June 2018 APP vintage we estimate that in the absence of the APP the 10-year sovereign bond yield would have been around 95 bps higher at that point (Figure 1).

5.2 Term Premium Compression over Time

Figure 9 plots the term premium impact at the 10-year maturity for different APP vintages over time. At each point in time indicated on the horizontal axis, the figure shows the estimated 10-year term premium compression for the different APP vintages.

Figure 9 is constructed as follows. For each of the trajectories shown there, the starting point is the initial impact, i.e., the 10-year maturity point in Figure 1. For each of the trajectories shown, the impact over time is then obtained by moving to the right along the corresponding free-float compression curve in Figure 4. To this end, we use the impact formula (16)—reproduced here for convenience: $dy_n(\mathcal{U}_{t+h}) = \gamma^{n\prime}\mathcal{U}_{t+h}$—by applying the impact vector γ^{10y} to the sequences of anticipated innovations $\mathcal{U}$ that start in the future at time $t+h$. For example, for the June 2018 vintage we estimate the 10-year term premium reduction in January 2025, by taking the segment of the violet free-float impact curve in Figure 4 that starts in January 2025 and ends in December 2034 as our sequence for the free-float reduction. The scalar product with the time-invariant impact factor vector (γ^{10y}) then delivers an estimated 10-year yield impact of around 35 bps in January 2025.

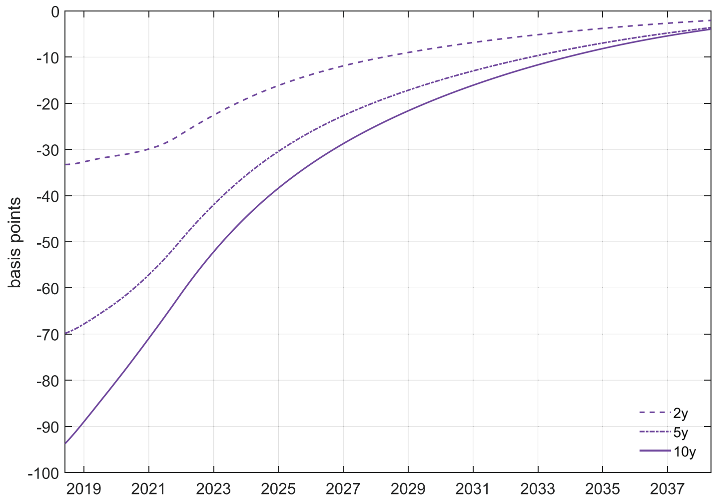
The estimated term premium impact is fairly persistent but gradually fades over time. Across the APP trajectories shown, the half-life of the initial impact on the 10-year yield is around five to six years. While the projected 10-year term premium compression falls below 10 bps by around 2033, it only dissipates completely once the portfolio has been entirely wound down.

For shorter maturities, the impact of the program also diminishes over time, albeit more slowly than at longer maturities; see Figure 10 for the June 2018 APP vintage. Looking at the 2-year maturity, the initial term premium effect is smaller than for the 10-year maturity, which implies a flattening of the curve, as discussed above. As the 2-year impact fades more slowly than the 10-year impact, the yield curve becomes again steeper over time. The markedly greater persistence of the two-year term premium compression over the nearer term reflects the impact of reinvestments, which were anticipated to follow the end of net purchases in December 2018 and assumed to last for three years. Hence, even in early 2020 most of the term of a two-year bond is falling into the reinvestment phase, which is not true for longer-term bonds.

The fading of the term premium compression reflects, to some extent, the "aging" of the portfolio—i.e., its gradual loss of duration as the securities held in the portfolio mature—as well as, in particular, the run-down of the portfolio that market participants anticipate will eventually follow the end of the expected horizon of reinvestments.

The pure "aging" effect is due to the fact that day by day the duration of the central bank portfolio falls even in the absence of any redemptions. The reinvestment of maturing principals conducted in

Figure 10. The Impact of the APP on the 2-Year, 5-Year, and 10-Year Term Premium over Time

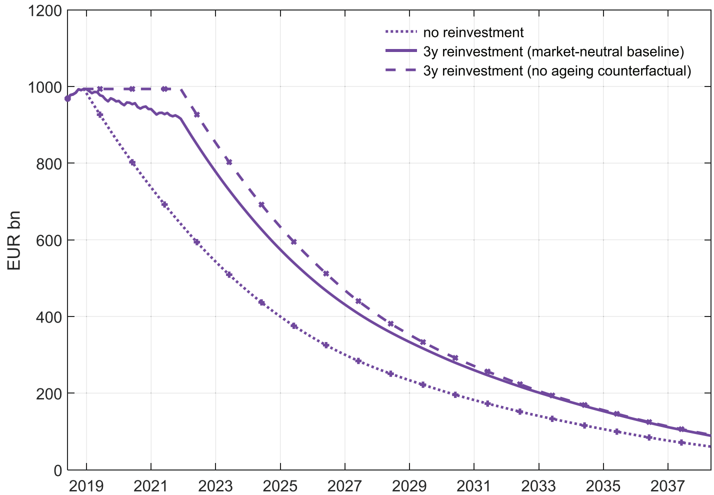


Note: The figure shows for the June 2018 APP vintage the impact of the APP on the term premium component of the 2-year, 5-year, and 10-year sovereign bond yield (averaged across the four largest euro area countries) over time.

line with "market neutrality"—i.e., with the maturity distribution of purchases guided by the maturity distribution of the eligible universe of securities—offsets this gradual loss of duration to some extent (see the continuous versus the dotted line in Figure 11) over the reinvestment horizon (assumed to be two years). By contrast, under a counterfactual "no aging" reinvestment policy (dashed-dotted line in Figure 11), the portfolio would remain constant in terms of 10-year equivalents during the reinvestment phase. Figure 12 illustrates the term premium compression that would result from such a counterfactual "no aging" reinvestment policy.[26] It turns out that even if the portfolio was prevented from aging during the reinvestment phase, the term premium impact of the central bank purchases would still fade gradually over time. This suggests that the bulk of the fading

[26]This type of reinvestment policy would be challenging to implement in practice, as it would require reinvestments into very long-term securities, with an average maturity of around 13 years.

Figure 11. Illustrating Portfolio Aging: The Evolution of Duration-Weighted Government Bond Holdings Under Different Reinvestment Scenarios

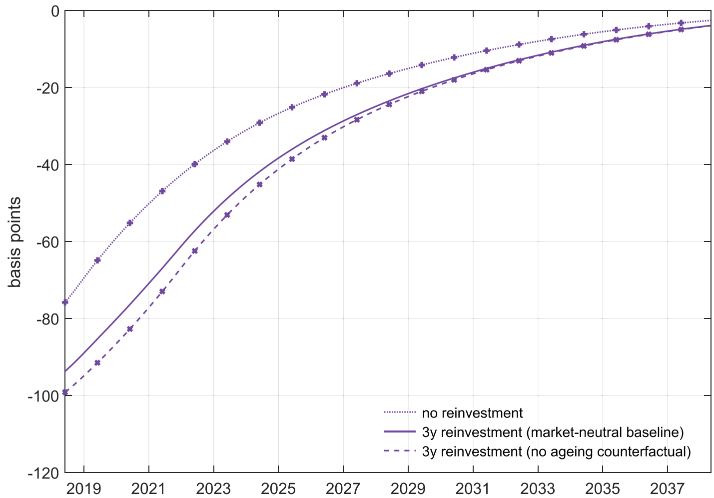


Note: For the June 2018 APP vintage of net purchases, the figure shows the projected evolution of the big-four government bond holdings in terms of 10-year equivalents under three reinvestment scenarios. Under the "no reinvestment" scenario, the portfolio starts running down after the end of net purchases in December 2018. In the "3y reinvestment scenario (market-neutral baseline)" scenario, reinvestments are made for three years starting in January 2019 in line with a "market-neutral" maturity distribution of purchases. In the "3y reinvestment (no ageing counterfactual)" scenario, reinvestments are again conducted over a period of three years starting in January 2019, but deviating from our baseline case it is assumed that reinvestments are made in sufficiently long maturities to offset the "aging" of the portfolio during the reinvestment phase.

term premium impact in the future reflects market expectations of a gradual roll-down of the portfolio after the end of reinvestments.

Apart from the relevance of reinvestment in mitigating the aging effect, the reinvestment horizon as such makes an important contribution to the reduction in term premia and its persistence over time. Figure 13 illustrates for the June 2018 APP vintage the 10-year term premium reduction for reinvestment horizons ranging from 0 to 10 years. The longer is the reinvestment horizon, the higher is the contemporaneous yield impact. However, the marginal impact of an additional year of reinvestment is shrinking with the length of the reinvestment horizon. For instance, reinvesting for 3 years instead

Figure 12. Illustrating Portfolio Aging: The APP's 10-Year Term Premium Impact under Different Reinvestment Scenarios



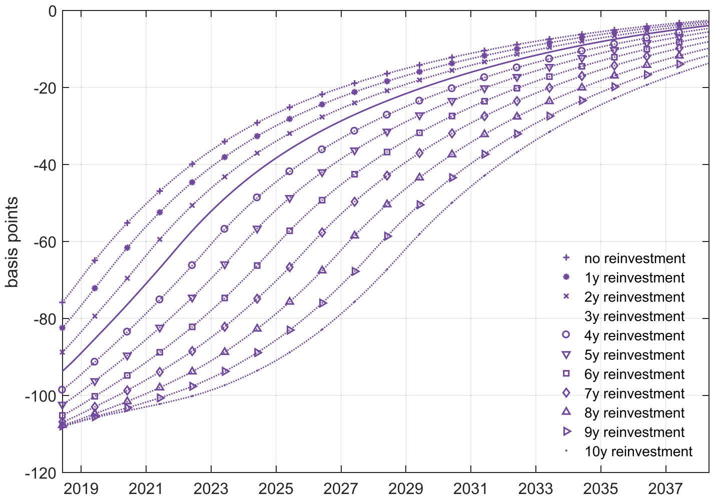
Note: The figure shows the 10-year term premium impact over time that is implied by the respective trajectory of central bank holdings shown in Figure 11.

of 0 years generates an additional term premium impact of around 18 bps, while moving from 7 to 10 years of reinvestment induces an additional compression of a mere 2 bps. This declining marginal effect reflects discounting: the additional free-float reduction in the 7-year versus 10-year reinvestment scenario happens 7 years from now, which is priced into contemporaneous term premia via low levels of impact factors; see Figure 7 again. But the picture changes over time: standing in, say, 2026, the marginal impact of going from 7-year to 10-year reinvestment (following the end of net purchases in December 2018) is larger than in June 2018.

5.3 Benchmarking Announcement Effects of APP Recalibrations

We benchmark the term premium impact of APP recalibration announcements predicted by our model against the observed yield curve changes within narrow windows around the respective announcement dates. Since our model estimation is not informed

Figure 13. Evolution of the 10-Year Term Premium Compression through the APP for Different Reinvestment Horizons



Note: For the June 2018 APP vintage, the figure shows the 10-year term premium compression over time for different reinvestment horizons following the end of net asset purchases in December 2018.

by post-March 2015 data, these exercises represent an out-of-sample cross-check of our model.

To calculate the surprise entailed by the APP recalibration announcements for the future free-float, we control for pre-announcement market expectations. For each APP recalibration we first simulate the free-float trajectory based on survey expectations about the path of the APP before the announcement and then again based on the actually announced purchase parameters (see Table 2). The difference between these two free-float trajectories gives us the sequence of surprises to the free-float due to the APP recalibration announcement. We feed these surprises into our model (using them as the $\mathcal{U}$ sequence in Equation (16)), and compare this model prediction with the one-day yield curve changes measured around the APP recalibration announcement date.

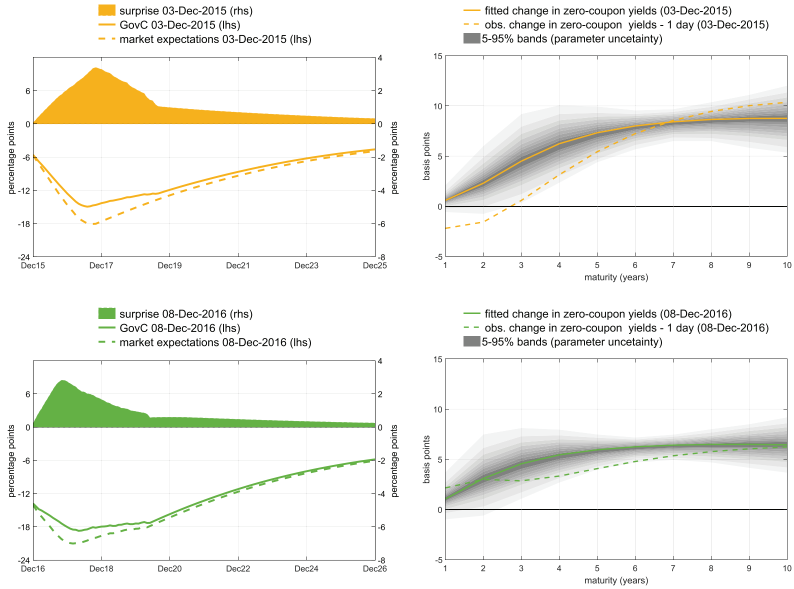
We cleanse the observed yield curve changes from both changes in the bond yield's expectations component (average short-rate expectations over the bond's maturity) as well as macro surprises.

This makes the yield changes more closely comparable to the yield changes predicted by our term structure model, which captures the change in yields purely based on the term premium compression via duration risk extraction. Specifically, first, to control for the expectations component we subtract from the full observed yield change the change in the estimated expectation component of the euro area swap (OIS) rate curve, which we obtain from a benchmark affine term structure model based on Joslin, Singleton, and Zhu (2011). Second, we account for macroeconomic surprises on the days of the announcements of APP recalibrations by cleansing yield changes for macro effects relying on the sensitivity of yields to macroeconomic surprises obtained in Section 4.1. The yield changes shown on the right-hand side of Figure 14 are the average of the observed yield changes cleansed of macro effects and those not cleansed of such effects (but in both cases cleansed of changes in short-rate expectations).

For the December 2015 and December 2016 APP recalibrations, Figure 14 shows on the left side the surprises in the free-float sequences and on the right side the corresponding model-implied changes in the yield curve, as well as the (cleansed) observed yield curve changes. Among the dates at which the ECB recalibrated its purchase program, the December 2015 and December 2016 announcements stand out as those that surprised the market the most and triggered the strongest yield response. Both recalibration announcements implied less duration extraction by the ECB than the market anticipated and therefore some upward revisions to the free-float of duration risk.[27] Yields increased for all but the

[27]In December 2015, the ECB announced the first recalibration of the APP since the initial announcement of the program in January 2015. The recalibration involved the announcement of, first, a prolongation of net purchases by six months until March 2017 at an unchanged purchase pace of €60 biliion per month, and, second, a reinvestment policy for maturing principals beyond the net purchase horizon "for as long as necessary." While market participants had anticipated the prolongation of net asset purchases ahead of the December 2015 Governing Council, the predominant expectation had been for the ECB to also increase the monthly pace of purchases (see Table 2). As a result, and despite the announcement of the reinvestment policy, the APP recalibration implied a significantly lower duration absorption over the near term than expected by market participants. The December 2016 recalibration featured an extension of net purchases at a reduced monthly pace of €60 billion for nine months until December 2017.

Figure 14. Announcement Effects of Selected APP Recalibrations



Note: Left-hand side: The dashed lines represent the expected APP-induced free-float compression paths constructed from surveys before the APP recalibration was announced. The continuous lines represent the expected free-float compression paths after the APP recalibration announcement, constructed based on the announced APP parameters. The difference between the pre- and post-APP recalibration free-float projections is the free-float surprise, which is shown as an area. The right-hand side shows the model-implied term premium impact due to the free-float surprises (APP recalibrations versus observed yield changes). The observed changes (cleansed from changes in short-rate expectations and macroeconomic surprises) are shown as dashed lines. The gray area is the bootstrap-based 5–95 percent confidence band of the model-implied yield changes reflecting parameter uncertainty.

shortest maturities on the day of the Governing Council announcement, thereby steepening the curve. Filtering the resulting revision

---

Market expectations were for an extension over a somewhat shorter horizon at a slightly higher monthly pace; see Table 2. In addition, the ECB expanded the eligible maturity bracket from two to one year at the lower end and also opened the door to purchases at yields below the deposit facility rate "to the extent necessary." The somewhat lower monthly purchase pace and increased scope for buying short-term papers implied some upward surprise on the expected path of the free-float.

of the free-float sequence through our model can broadly explain the observed reaction of the yield curve: the model predicts the right shape of change, i.e., the curve steepening, and (cleansed) observed yield changes are overall covered by our bootstrap-based 5–95 percent confidence intervals for model-implied yield changes.[28]

By contrast, the remaining APP recalibrations were closely in line with market expectations. As a result, for those recalibrations there was no revision to the expected free-float of duration risk that helps identify how well the model captures revisions of the free-float sequence. In the absence of such revisions, our model predicts no changes in yields for those APP recalibrations. On those days, nonetheless, some mild movements in the yield curve were observed, which likely reflect changes in other policy instruments—in particular, the cut of the deposit facility rate that accompanied the March 2016 APP recalibration and the change in the ECB's forward guidance on the path of policy rates that came with the June 2018 APP recalibration.

5.4 Uncertainty and Robustness of Results

We conduct sensitivity analyses around our baseline results via four avenues: first, we account for parameter uncertainty based on bootstrapping; second, we conduct a bias adjustment of the estimated factor dynamics; third, we vary our estimation sample; fourth, we reestimate the model with a differently scaled bond supply variable.

To account for parameter uncertainty, we rely on a bootstrap procedure. We do so, as our two-step estimation approach and the limited number of available observations prevent a straightforward application of asymptotic results. For the bootstrap we resample the data and obtain bootstrap estimates of the model parameters based on our two-step estimation approach outlined in Section 4.1.

In detail, in the ith bootstrap run, for the first step of our estimation approach, we take random draws from the centered residuals of our estimated factor VAR and use them as innovations in generating a new time series of factors, based on the point estimates

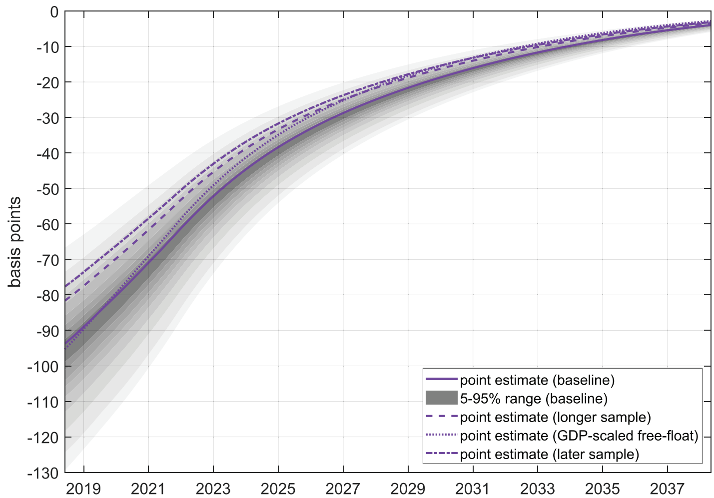
[28]See Section 5.4 for details on the construction of the confidence intervals.

of VAR parameters $\hat{c}$ and $\hat{\mathcal{K}}$. Using those bootstrap factor realizations, we reestimate the factor VAR parameters—under the same zero restrictions as in our baseline estimation—and obtain bootstrap estimates $\check{c}^{(i)}$, $\check{\mathcal{K}}^{(i)}$, and $\check{\Omega}^{(i)}$. Similarly, we reestimate the short-rate equation (2) and obtain bootstrap estimates $\check{\delta}_0^{(i)}$ and $\check{\delta}_1^{(i)}$ of δ_0 and δ_1. For the second step of our estimation, we bootstrap the realizations of the components featuring in our dual objective function (26). For the first part of that objective function, F_1, we construct a bootstrap realization of the time series of yields by adding measurement errors to fitted yields, where these errors are drawn from the pool of centered fitting residuals of our estimated model. For the second component, F_2, we then generate a bootstrap realization of the change in the yield curve over our event window by applying noise around the fitted yield changes.[29] Given the bootstrap draw of the yield changes over the event window and the bootstrap draw of the yields sequence, we conduct the second step of our estimation approach, i.e., we minimize the dual objective criterion (26) for given $\check{c}^{(i)}$, $\check{\mathcal{K}}^{(i)}$, $\check{\Omega}^{(i)}$, $\check{\delta}_0^{(i)}$, and $\check{\delta}_1^{(i)}$ and obtain bootstrap estimates $\check{\lambda}_0^{(i)}$ and $\check{\Lambda}_1^{(i)}$ of λ_0 and Λ_1, respectively. We repeat this procedure for $K = 1{,}000$ bootstrap repetitions. Collecting our parameters in a vector θ, the distribution of our point estimate $\hat{\theta}$ is hence approximated by the sampling distribution $(\check{\theta}^{(1)}, \ldots, \check{\theta}^{(K)})$ of our bootstrap estimates. Similarly, the distribution of (non-linear) functions of the parameters $g(\theta)$—like, e.g., the impact factors $\gamma_h^n \equiv \gamma_h^n(\theta)$—are approximated by the corresponding bootstrap sampling distribution $g(\check{\theta}^{(1)}), \ldots, g(\check{\theta}^{(K)})$. This enables us to generate distributions around all our impact estimates that reflect the uncertainty stemming from parameter estimation.

Figure 15 shows the APP's dynamic impact on the 10-year term premium based on the June 2018 APP vintage with our estimated confidence bands. The midpoint (solid violet line) is the same as in Figure 9. The uncertainty band is the bootstrap-based confidence band reflecting parameter uncertainty. For the contemporaneous

[29] Proceeding in the standard fashion as we did for the bootstrap generation of factors and yields, we could draw from the residuals corresponding to the fit of our six yield changes over the event window. However, as these residuals are very small (not exceeding 2 bps), we see a risk of underestimating the true uncertainty and take a more conservative approach by drawing the noise from six independent normal distributions with zero mean and a standard deviation of 10 bps.

Figure 15. The APP's 10-Year Term Premium Impact: Parameter Uncertainty and Sample Robustness



Note: Conditional on June 2018 information, the figure shows the impact of the APP on the 10-year term premium over time. The solid line is the point estimate, identical to the violet line in Figure 9. The shaded area is the 5–95 percent confidence band stemming from parameter uncertainty computed using a bootstrap approach. The dotted line ("GDP-scaled free-float") corresponds to a model specification that uses the nominal GDP of the big-four euro area countries for scaling the free-float ratio instead of overall bond supply. The dashed line ("longer sample") is an alternative point estimate that uses the full-sample yield information (until June 2018) for estimation, whereas the baseline estimate ignores data after March 2015. The dashed-dotted ("later sample") line corresponds to an estimation that uses data as of September 2014 only.

term premium impact as of June 2018, the 5–95 percent confidence band ranges from 65 to 130 bps around the 95 bps estimate. The width of the confidence bands accounting for parameter uncertainty is of the same order of magnitude as that reported in Ihrig et al. (2018). Over time the uncertainty band gradually narrows, as the point estimate and the uncertainty around it converges to zero. Formally, this can be seen from the fact that at any future point in time $t + h$ the yield impact is given by the product of impact factors and a sequence of APP free-float innovations going forward, $dy_n(\mathcal{U}_{t+h}) = \sum_{k=1}^{n} \gamma_k^n(\theta) u_{t+h+k-1}$: as the innovations u_{t+h+k} eventually shrink to zero, so does the overall scalar product.

As a second robustness check, we conduct a bias adjustment of the estimated factor dynamics. As noted in the literature,[30] term structure models tend to underestimate the high persistence exhibited by bond yields—in particular, when the estimation sample is short. Bauer, Rudebusch, and Wu (2012) and others have, therefore, suggested to conduct a bias correction when estimating the VAR dynamics of factors. For our model and data, in fact, the estimated degree of persistence of the factor VAR is already high, with a maximum eigenvalue of 0.976 (Table 3). Nevertheless, we apply the bias correction methods suggested by Bauer, Rudebusch, and Wu (2012).[31] Overall, the bias correction leaves the main results regarding the APP's impact on the yield curve essentially unchanged. We attribute this to the fact that, despite different dynamics of factors under the $\mathbb{P}$ measure, the cross-section information in bond yields, especially their change over the event-window dates, ensures that key objects like the impact factors γ_h^n in (16), which depend on both $\mathbb{P}$ and market-price-of-risk parameters, are hardly affected.

Thirdly, we vary the estimation sample. In our baseline specification we only use data up to March 2015 to estimate the model. This approach allows undertaking a clean out-of-sample benchmarking exercise over the period of the Eurosystem asset purchases, as discussed in Section 5.3. One robustness check we conducted is to use the full data set spanning December 2009 to June 2018 to estimate the model. Specifically, for the first step of the estimation approach we estimate the factor VAR and the short-rate equation over the full sample. For the second step, we leave the second component of the dual objective function in Equation (26) unchanged, but include the full time series of bond yields until June 2018 in the first component of the dual objective function. As shown in Figure 15, dashed line, using the full sample, the estimated impact of the APP is of a similar magnitude, if somewhat smaller (81 versus 93 bps on impact in June 2018; and 42 versus 48 bps after five years). Furthermore, we conducted an additional robustness check by estimating the model using only data as of the time asset purchases started being anticipated,

[30]See Bauer, Rudebusch, and Wu (2012) and Kim and Orphanides (2012).

[31]We deploy an analytical bias approximation, a bootstrap-based bias correction, and an indirect inference estimator for bias correction, all based on the code for Bauer, Rudebusch, and Wu (2012) provided on Cynthia Wu's website.

i.e., September 2014, until June 2018. In order to ensure comparability with the baseline estimation, we left the data used for the event-study featuring in the first part of the dual objective function—September 2014 to March 2015—unchanged. The strength of the yield impact estimates (dashed-dotted line) decreases further under this setup, but only marginally compared with the one using the full data set, and it remains within the range of the bootstrap confidence bands associated with the baseline approach.

As a final robustness check, we rerun the model estimation and selected policy analysis with a different free-float measure that uses the total nominal GDP of the big-four euro area countries as scaling variable in the denominator instead of total bond supply as in our baseline scarification. Doing so, we estimate a broadly similar yield impact (see the dotted line in Figure 15). This reflects the fact that the two free-float measures are highly correlated. Moreover, the reestimated parameters obtained with the alternative free-float measures partially compensate for the differences in the free-float measures.

6. Conclusion

Central bank bond purchases extract duration risk which otherwise would be borne by arbitrageurs. This decreases the market price of risk and compresses the term premium component of bond yields across the term structure. Our paper quantifies the strength of the duration channel for the European Central Bank's asset purchase program (APP). We deploy an arbitrage-free term structure model along the lines of Li and Wei (2013). In addition to the level and slope factors, aggregate duration affects the market price of level and slope risk and hence the term premium across maturities. This link between bond supply and yields reflects one of the key features of the microfounded equilibrium model of Vayanos and Vila (2021).

We find that, first, the contemporaneous impact of the ECB's APP flattens the yield curve and amounts to around 95 bps by mid-2018 for the 10-year maturity. This impact is comparable to point estimates found for the Federal Reserve's large-scale asset purchase programs. Second, the effect is persistent and expected to only slowly fade over time, with a half-life of around five years. Third, the expected length of the reinvestment period after net purchases has a significant impact on term premia. For example, as of June

2018, relative to a counterfactual of no reinvestment, an expected reinvestment horizon of three years compressed term premia by an additional 18 bps. Finally, recalibrations of APP purchase parameters imply surprises for the central bank's expected path of duration extraction. Overall, our model accounts well—in real time—for the duration-implied yield curve impact of such recalibrations on the term structure of interest rates, while at the same time other factors—going beyond the duration channel—can move the yield curve around such announcements but are outside the scope of our model.

Our contribution to the literature is threefold. First and foremost, our paper is the first to provide a comprehensive assessment of the contemporaneous and dynamic effects of the ECB's APP across the term structure and their evolution over time. By contrast, other available studies have largely focused on the impact of the initial announcement impact of the APP on asset prices. Second, based on security-level information of asset holdings, aggregate issuance data, and ECB portfolio holdings, we construct a novel granular measure of the "free-float of duration risk," i.e., the duration-weighted share of public sector debt in the hands of arbitrageurs. Accordingly, our measure tries to mirror the theory set out in Vayanos and Vila (2021). Moreover, we construct projections of that free-float measure, which is a crucial input for the model-based translation of changes in APP purchase parameters into changes in the term premium. To this end, we not only rely on the purchase parameters announced by the ECB, but also account for market expectations by exploiting survey expectations on the path of ECB asset purchases and projecting the market-expected trajectory of reductions in the free-float due to the APP. Third, on the methodological side, we meet the constraints imposed by the relatively short time series of euro area data by deploying a new two-step estimation approach that relies on both fitting the time series of bond yields as well as on utilizing event-based information in the run-up to the ECB's APP. Given this non-standard approach, we also provide a bootstrap procedure to gauge the impact of parameter uncertainty on our estimates.

While our approach utilizes a reduced-form term structure model incorporating the no-arbitrage condition and a stylized version of the duration extraction channel formulated in Vayanos and Vila (2021), our analysis can inform a more structural modeling of the duration channel of central bank asset purchases. In particular, it could help support the specification and quantification of microfounded equilibrium models.[32]

We acknowledge that several issues relating to modeling, measurement, and the scope of analysis remain to be addressed in future work.

On the modeling side, first, whereas Li and Wei (2013) and this paper serve as examples of how intricate it is to capture the stock effect of QE by itself, there is a case for developing more encompassing models that allow to also incorporate signaling and flow effects in a unified framework. Second, our linear model does not incorporate a lower bound on interest rates. The proximity of rates to the lower bound is potentially relevant for the impact of QE, as documented in King (2019). However, incorporating the lower bound into a term structure model with a duration channel is challenging, especially if there is a need to account for a negative and time-varying level of the lower bound, as would be warranted in the euro area context. Third, another dimension for further refinement is distinguishing the effect QE has on real term premia versus inflation risk premia. While our model captures the impact of QE on the overall nominal term premium, the transmission of QE via the duration channel is likely to affect bond yields mainly via real term premia; see also Kaminska and Zinna (2020) and the references therein. Fourth, it would also be useful—but equally difficult in our specific model framework—to allow for relaxing the assumption of zero risk compensation for supply uncertainty and to build a model that can account for time variation in supply uncertainty in order to capture taper-tantrum type bouts of higher uncertainty. Fifth, focusing in more detail on the implementation of purchases, enhanced models may take into account the impact of purchases on both bond risk premia and liquidity premia.[33]

[32]For example, King (2015) examines features that are necessary—in particular, with regard to the properties of the implied stochastic discount factor—for general equilibrium models to exhibit a duration channel of the kind we analyze in this paper.

[33]See, e.g., the calibrated search-theoretic model by Ferdinandusse, Freier, and Ristiniemi (2020) as regards the impact of central bank purchases on market liquidity and the relevance of preferred-habitat investors for this channel.

With regard to the measurement duration risk, further work to examine the robustness of the supply variable is warranted. This includes both the analysis of the portfolio rebalancing behavior of different investors and their appropriate assignment to the group of arbitrageurs, and the normalization of the supply variable. With regard to the latter, it would be useful to examine ways to account more directly for the "risk-bearing" capacity of arbitrageurs, for which some measure of investors' size or capital are worth exploring as alternatives to the total bond supply or GDP.

As for the scope of analysis, our approach could be taken further by also studying the impact of the APP on a wider range of asset classes—in particular, corporate bonds—or by taking a more disaggregated view on the impact across individual euro area countries.[34] Finally, the ECB (and other central banks) have expanded their asset purchases in response to the economic downturn induced by the COVID-19 pandemic. Studying the effect of purchases during that period in an environment of higher risk aversion may grant further insights into how they are transmitted to the yield curve.

Appendix. Comparing the Empirical Model to Vayanos and Vila (2021)

The reduced-form empirical model used in this paper and the U.S. counterpart by Li and Wei (2013) are inspired by the equilibrium model by Vayanos and Vila (2021).[35] As a key feature common to both models, the outstanding supply of debt that has to be borne by arbitrageurs affects the market price of risk of factor exposure and hence term premia. The commonality is best seen when focusing on the model version in Vayanos and Vila (2021), Section 3.2, that switches off time variation in the demand of preferred-habitat investors. Bonds are in zero net supply, and maturity-specific preferred-habitat investor demand is constant, so that $x^{(\tau)}$ has to be held by arbitrageurs. In the one-factor version of

[34] See recent work by Costain, Nuño, and Thomas (2022), who present a calibrated model incorporating both a duration extraction and a sovereign credit risk channel to quantify the impact of the ECB's asset purchases during the pandemic phase on individual sovereign yields.

[35] See also Greenwood and Vayanos (2014).

their model, with short-rate exposure as the only risk, the (instantaneous) expected excess return of holding the τ-period bond at time t is given by $\mathcal{A}_i(\tau)\lambda_{i,t}$, where $\mathcal{A}_i(\tau)$ is the loading of the log bond price on the short-rate factor.[36] The market price of short-rate risk is given by

$$\lambda_{i,t} = a\sigma_i^2 \int_0^T x_t^{(\tau)} \mathcal{A}_i(\tau) d\tau, \tag{27}$$

where a is the risk-aversion parameter, σ_i^2 is the innovation volatility to the short-rate process, $x_t^{(\tau)}$ is outstanding bond supply of maturity τ (i.e., whatever is left net of preferred-habitat holdings) to be absorbed by the arbitrageur, and T is the maximum maturity of outstanding debt.

In the empirical model used here, the short-term rate is driven by two factors (level and slope). The market price of level/slope risk is given in (20). The time-varying contribution of the supply factor to the factor price is $\Lambda_{1,ZQ} \cdot Q_t$, where Z denotes either level or slope and Q is our supply variable. This expression is of the same form as in (27), i.e., it is a product of a constant coefficient ($a\sigma_i^2$ versus $\Lambda_{1,ZQ}$) and a quantity variable ($\int_0^T x_t^{(\tau)} \mathcal{A}_i(\tau) d\tau$ versus Q_t).

Regarding the time-constant multiplier, our reduced-form parameter $\Lambda_{1,ZQ}$ may hence be interpreted as reflecting risk aversion. At the same time, though, it has to be noted that such a mapping from a structural to a reduced-form model (with more factors) is necessarily incomplete.

Regarding the supply variable, the expression $\int_0^T x_t^{(\tau)} \mathcal{A}_i(\tau) d\tau$ in Equation (27) can be interpreted as aggregate duration risk in the market. $\mathcal{A}_i(\tau)$ is the individual bond's (with maturity τ) exposure to short-rate risk. This is weighted by the outstanding bond supply $x_t^{(\tau)}$ for that maturity and summed up (integrated) across maturities. Greenwood and Vayanos (2014) compare that measure to "simple dollar duration" defined as $\int_0^T x_t^{(\tau)} \tau d\tau$, i.e., the weighting is not the bond-specific sensitivity but simply the maturity of the respective bond. Note that the latter expression is analogous to that appearing in the enumerator of our free-float measure Q in (1), i.e.,

[36] We slightly adapt their notation to avoid overlaps with the symbols used in this paper.

multiplying maturities with corresponding supply volumes prevailing in those maturity brackets. For Greenwood and Vayanos (2014), this measure of simple dollar duration turns out to be closely correlated to their model-implied (using their parameter calibration) counterparts of short-rate and supply-duration risk.

Summing up, the overall economic mechanism through which quantitative easing affects the term premium is the same in both the equilibrium model of Vayanos and Vila (2021) and the non-structural empirical model used here: an increase in future expected central bank purchases would reduce (current and) expected aggregate duration risk to be absorbed by the market. This reduces the market price of risk, which leads to lower expected excess returns in the future and hence to a contemporaneous compression of term premia and bond yields across maturities. While the equilibrium model and the empirical models share the relevance of bond quantities as a key property, they differ in other details. In particular, their full model (as opposed to the simplified version explained above) features several demand factors determining variation in the demand of preferred-habitat investors. This modeling choice renders the bond volume to be held by the arbitrageur a multi-factor process and these demand factors are also priced in equilibrium. Adopting a factor structure for bond quantities and allowing for pricing of such factor risk is certainly high on the agenda for developing our empirical model further.

References

Adrian, T., R. Crump, and E. Moench. 2013. "Pricing the Term Structure with Linear Regressions." *Journal of Financial Economics* 110 (1): 110–38.

Altavilla, C., G. Carboni, and R. Motto. 2021. "Asset Purchase Programs and Financial Markets: Lessons from the Euro Area." *International Journal of Central Banking* 17 (4): 1–48.

Andrade, P., J. Breckenfelder, F. De Fiore, P. Karadi, and O. Tristani. 2016. "The ECB's Asset Purchase Programme: An Early Assessment." ECB Working Paper No. 1956.

Arrata, W., and B. Nguyen. 2017. "Price Impact of Bond Supply Shocks: Evidence from the Eurosystem's Asset Purchase Program." Working Paper No. 623, Banque de France.

Bauer, M. D., G. D. Rudebusch, and J. C. Wu. 2012. "Correcting Estimation Bias in Dynamic Term Structure Models." *Journal of Business and Economic Statistics* 30 (3): 454–67.

Bergant, K., M. Fidora, and M. Schmitz. 2020. "International Capital Flows at the Security Level: Evidence from the ECB's Asset Purchase Programme." ECB Working Paper No. 2388.

Blattner, T. S., and M. A. S. Joyce. 2020. "The Euro Area Bond Free Float and the Implications for QE." *Journal of Money, Credit and Banking* 52 (6): 1361–95.

Bouabdallah, O., C. Checherita-Westphal, T. Warmedinger, R. de Stefani, F. Drudi, R. Setzer, and A. Westphal. 2017. "Debt Sustainability Analysis for Euro Area Sovereigns: A Methodological Framework." ECB Occasional Paper No. 185.

Bua, G., and P. G. Dunne. 2019. "The Portfolio Rebalancing Effects of the ECB's Asset Purchase Programme." *International Journal of Central Banking* 15 (5): 1–46.

Bulligan, G., and D. Delle Monache. 2018. "Financial Markets Effects of ECB Unconventional Monetary Policy Announcements." Occasional Paper No. 424, Bank of Italy.

Christensen, J. H. E., and S. Krogstrup. 2019. "Transmission of Quantitative Easing: The Role of Central Bank Reserves." *Economic Journal* 129 (617): 249–72.

Costain, J., G. Nuño, and C. Thomas. 2022. "The Term Structure of Interest Rates in a Heterogeneous Monetary Union." Documento de Trabajo 2223, Banco de España.

D'Amico, S., W. English, D. López-Salido, and E. Nelson. 2012. "The Federal Reserve's Large-Scale Asset Purchase Programmes: Rationale and Effects." *Economic Journal* 122 (564): F415–F446.

D'Amico, S., and T. B. King. 2013. "Flow and Stock Effects of Large-Scale Treasury Purchases: Evidence on the Importance of Local Supply." *Journal of Financial Economics* 108 (2): 425–48.

De Pooter, M., R. Martin, and S. Pruitt. 2018. "The Liquidity Effects of Official Bond Market Intervention." *Journal of Financial and Quantitative Analysis* 53 (1): 243–68.

De Santis, R. A. 2020. "Impact of the Asset Purchase Programme on Euro Area Government Bond Yields Using Market News." *Economic Modelling* 86 (March): 192–209.

De Santis, R. A., and F. Holm-Hadulla. 2020. "Flow Effects of Central Bank Asset Purchases on Sovereign Bond Prices: Evidence from a Natural Experiment." *Journal of Money, Credit and Banking* 52 (6): 1467–91.

Eser, F., and B. Schwaab. 2016. "Evaluating the Impact of Unconventional Monetary Policy Measures: Empirical Evidence from the ECB's Securities Markets Programme." *Journal of Financial Economics* 119 (1): 147–67.

Ferdinandusse, M., M. Freier, and A. Ristiniemi. 2020. "Quantitative Easing and the Price-Liquidity Trade-off." ECB Working Paper No. 2399.

Geiger, F., and F. Schupp. 2018. "With a Little Help from My Friends: Survey-Based Derivation of Euro Area Short Rate Expectations at the Effective Lower Bound." Discussion Paper No. 27, Deutsche Bundesbank.

Ghysels, E., J. Idier, S. Manganelli, and O. Vergote. 2016. "A High-Frequency Assessment of the ECB Securities Markets Programme." *Journal of the European Economic Association* 15 (1): 218–43.

Greenwood, R., and D. Vayanos. 2014. "Bond Supply and Excess Bond Returns." *Review of Financial Studies* 27 (3): 663–713.

Hamilton, J. D., and J. C. Wu. 2012. "The Effectiveness of Alternative Monetary Policy Tools in a Zero Lower Bound Environment." *Journal of Money, Credit and Banking* 44 (1): 3–46.

Hartmann, P., and F. Smets. 2018. "The European Central Bank's Monetary Policy during Its First 20 Years." *Brookings Papers on Economic Activity* (2): 1–146.

Ihrig, J., E. Klee, C. Li, M. Wei, and J. Kachovec. 2018. "Expectations about the Federal Reserve's Balance Sheet and the Term Structure of Interest Rates." *International Journal of Central Banking* 14 (2): 341–90.

Joslin, S., K. J. Singleton, and H. Zhu. 2011. "A New Perspective on Gaussian Dynamic Term Structure Models." *Review of Financial Studies* 24 (3): 926–70.

Joyce, M. A. S., and M. Tong. 2012. "QE and the Gilt Market: A Disaggregated Analysis." *Economic Journal* 122 (564): F348–F384.

Kaminska, I., and G. Zinna. 2020. "Official Demand for U.S. Debt: Implications for U.S. Real Rates." *Journal of Money, Credit and Banking* 52 (2–3): 323–64.

Kandrac, J., and B. Schlusche. 2013. "Flow Effects of Large-Scale Asset Purchases." *Economics Letters* 121 (2): 330–35.

Kim, D. H., and A. Orphanides. 2012. "Term Structure Estimation with Survey Data on Interest Rate Forecasts." *Journal of Financial and Quantitative Analysis* 47 (1): 241–72.

Kim, D. H., and J. H. Wright. 2005. "An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates." FEDS Working Paper No. 2005-33, Board of Governors of the Federal Reserve System.

King, T. B. 2015. "A Portfolio-Balance Approach to the Nominal Term Structure." Working Paper No. 2013-18, Federal Reserve Bank of Chicago.

———. 2019. "Expectation and Duration at the Effective Lower Bound." *Journal of Financial Economics* 134 (3): 736–60.

Koijen, R. S. J., F. Koulischer, B. Nguyen, and M. Yogo. 2017. "Euro-Area Quantitative Easing and Portfolio Rebalancing." *American Economic Review* 107 (5): 621–27.

———. 2021. "Inspecting the Mechanism of Quantitative Easing in the Euro Area." *Journal of Financial Economics* 140 (1): 1–20.

Lemke, W., and T. Werner. 2020. "Dissecting Long-Term Bund Yields in the Run-up to the ECB's Public Sector Purchase Programme." *Journal of Banking and Finance* 111 (February): Article 105682.

Li, C., and M. Wei. 2013. "Term Structure Modeling with Supply Factors and the Federal Reserve's Large-Scale Asset Purchase Programs." *International Journal of Central Banking* 9 (1): 3–39.

Rostagno, M., C. Altavilla, G. Carboni, W. Lemke, R. Motto, and A. S. Guilhem. 2021. "Combining Negative Rates, Forward Guidance and Asset Purchases: Identification and Impacts of the ECB's Unconventional Policies." ECB Working Paper No. 2564.

Schlepper, K., H. Hofer, R. Riordan, and A. Schrimpf. 2017. "Scarcity Effects of QE: A Transaction-Level Analysis in the Bund Market." BIS Working Paper No. 625.

Vayanos, D., and J. Vila. 2021. "A Preferred-Habitat Model of the Term Structure of Interest Rates." *Econometrica* 89 (1): 77–112.

A Comparison of Fed "Tightening" Episodes since the 1980s*

Kevin L. Kliesen
Federal Reserve Bank of St. Louis

This article examines how the real economy and inflation and inflation expectations evolved in response to the six tightening episodes enacted by the FOMC from 1983 to 2018. The findings indicate that the sixth episode (2015–18) differed in several key dimensions compared with the previous five episodes. In the first five episodes, the data show the FOMC was generally tightening into a strengthening economy with building price pressures. In contrast, in the final episode the FOMC began its tightening regime during a deceleration in economic activity and with headline and core inflation remaining well below the FOMC's 2 percent inflation target. Moreover, both short- and long-term inflation expectations were drifting lower. These developments helped explain why there was a one-year gap between the first and second increases in the federal funds target rate in the final episode. Another key difference is that in three of the first five episodes, the FOMC continued to tighten after the yield curve inverted; a recession then followed shortly thereafter. However, in the final episode, the FOMC ended its tightening policy about eight months before the yield curve inverted.

JEL Codes: E3, E4, E5, N1.

---

*Kevin L. Kliesen is a research officer and economist at the Federal Reserve Bank of St. Louis (https://research.stlouisfed.org/econ/kliesen/sel/). The author thanks the editor, an anonymous referee, and Michael McCracken for helpful comments. Any errors are the responsibility of the author. Kathryn Bokun, Rachel Harrington, Brian Levine, Lowell Ricketts, and Devin Warner provided excellent research assistance. Views expressed not necessarily those of the Federal Reserve Bank of St. Louis or the Board of Governors of the Federal Reserve System. Author e-mail: Kevin.L.Kliesen@stls.frb.org.

> "The FOMC has always recognized that in a tightening cycle, if we stop too soon, inflationary pressures will resurge and make it very difficult to contain them again. We therefore always tend to take out the insurance of an additional fed funds increase, fully expecting that it may not be necessary."
>
> Former Federal Reserve Chairman Alan Greenspan[1]

1. Introduction

The Federal Open Market Committee (FOMC) voted to establish a target range for the intended federal funds rate of 0 percent to 0.25 percent at the conclusion of its December 15–16, 2008, meeting. Although the decision was implemented during one of the nation's worst economic and financial crises, this decision was nonetheless historic. In the FOMC's Greenbook prepared for this meeting, Board staff predicted that the federal funds target rate would remain at the zero (effective) lower bound (ZLB) through the end of 2012.[2] But this four-year period turned out to be an unprecedented seven years.

With the economy into the sixth year of expansion, and inflation pressures projected to increase modestly, the FOMC announced at the conclusion of its December 15, 2015, meeting that it was raising its target range by 25 basis points. (Note: Henceforth, the analysis will characterize the midpoint of the range as the federal funds target rate, or FFTR.) The initial tightening action—defined as the first increase in the FFTR during the sequence of increases—was the first since June 2006.

Following liftoff in December 2015, the FOMC then paused for a year. Not only was inflation well below the 2 percent target rate at liftoff, but low inflation also persisted through 2016. Moreover, short-term inflation expectations also drifted lower, while long-term inflation expectations remained unchanged at 2 percent. From this standpoint, as will be discussed in this paper, economic conditions

[1] Greenspan (2007, p. 156).

[2] The Greenbook was the economic conditions and forecast document distributed by the Board staff before each FOMC meeting. It is now known as the Tealbook Part A. See the Long-Term Outlook table on pages I–18 in the December 10, 2008, Greenbook: https://www.federalreserve.gov/monetarypolicy/files/FOMC20081216gbpt120081210.pdf.

at the initial stages of the 2015 episode were unique compared with previous episodes. In their defense, the FOMC announced that even after this tightening action, monetary policy was accommodative.

Altogether, from December 2015 to December 2018, the FOMC lifted its policy rate nine times, in increments of 25 basis points. At the conclusion of the December 18–19, 2018, meeting, the FOMC's target range for the federal funds rate was 2.25 to 2.5 percent. This turned out to be the final increase in this tightening episode, and the sixth to have occurred during the Great Moderation (post-1983) period.[3] With downside risks to the economy emerging, the FOMC reduced its policy rate by 75 basis points over the second half of 2019. The FOMC then returned the policy rate to the ZLB in mid-March 2020 because of the contraction in economic activity spawned by COVID-19. On June 8, 2020, the National Bureau of Economic Research Business Cycle Dating Committee announced that the nation's longest business expansion ended in February 2020.

This article will proceed as follows. Section 2 will identify the six tightening episodes during this period and briefly discuss economic and financial conditions during each episode. Sections 3 and 4 will discuss the Board staff's forecast accuracy for real GDP growth before, during, and after the tightening episodes, and then whether financial market participants accurately gauged the extent of the tightening at the beginning of each episode. Section 5 examines responses of key economic and inflation measures before and after each episode. Section 6 concludes.

2. Six Tightening Episodes

The literature that examines economic and financial market developments during individual U.S. monetary tightening episodes is relatively sparse. For this paper, a tightening action occurs when the FOMC votes to raise the FFTR.[4] Laforte and Roberts (2014) employ

---

[3]This article uses the term "tightening" to refer to policy decisions by the FOMC to raise the federal funds rate target.

[4]Most central bank models presume that raising the short-term nominal policy rate will, via the expectations effect, also raise key longer-term interest rates faced by firms, households, and governments who borrow in capital markets. Economic textbooks generally assert that business capital spending (fixed investment) is sensitive to changes in interest rates. However, the empirical evidence is less

the model used by the staff of the Board of Governors of the Federal Reserve System to show that a 100 basis point increase in the federal funds target increases the size of the output gap (real GDP as a percent of real potential GDP) by about 0.4 percentage point in the first two years, while lowering the core inflation rate by less than 0.1 percentage point.[5] Willems (2020) uses a data set of annual observations for 162 countries to show that an increase of 100 basis points in the central bank's target rate reduces real GDP by an average of 0.5 percent over a four- to five-year period. He finds the effect is nearly four times larger for advanced economies (−1.1 percent) than for emerging and developed economies (−0.3 percent).

Adrian and Estrella (2009) examined whether tightening cycles since 1955 helped to predict future economic outcomes. They concluded that most tightening actions generated increases in the unemployment rate and a narrowing or inversion of the yield curve. The latter, they argue, is a useful indicator of future economic activity during periods of tighter monetary policy. More recently, Orphanides (2015) examined episodes during the Great Moderation within the context of the pending normalization of monetary policy in 2015. Orphanides argued that the FOMC could improve economic outcomes by employing a more systematic policy (i.e., rules based) rather than a discretionary policy. Other contributions that examined policy discussions of past tightening episodes within a broader context (i.e., not a systematic analysis of individual episodes) can be found in Greenspan (2007) and Hetzel (2008).

This paper uses two primary criteria to identify beginning and ending dates of tightening episodes: FOMC documents (e.g., Records of Policy Actions or FOMC Statements issued after the meeting) and the time series of the FOMC's federal funds target rate to identify the daily dates of the beginning and ending of tightening episodes. The latter is useful because in the early 1980s the Committee more

supportive of this view. See Sharpe and Suarez (2015) for a recent assessment. More broadly, Willis and Cao (2015) use a time-series model to show that employment across most industries has become less sensitive to changes in the federal funds rate since 1984.

[5]The workhorse model is known as FRB/US. See http://www.federalreserve.gov/econresdata/notes/feds-notes/2014/november-2014-update-of-the-frbus-model-20141121.html. As the authors of this note show, other outcomes are possible if one makes different assumptions.

closely monitored growth of the M1 and M2 monetary aggregates. Moreover, they did not publicly announce when the federal funds rate was changed. It was not until the press release following the February 4, 1994, meeting that the FOMC begin to publicly communicate decisions to change the federal funds target rate in real time.

Table 1 lists the six tightening episodes based on these criteria: from March 1983 to August 1984; from March 1988 to May 1989; from February 1994 to February 1995; from June 1999 to May 2000; from June 2004 to June 2006; and from December 2015 to December 2018. A stricter definition of when a tightening cycle *ends* was developed by Adrian and Estrella (2009). They define the end of a cycle based on a set of criteria for the level of the federal funds rate relative to the beginning of the cycle or the end of the cycle. Comparing the ending points in Table 1 with Adrian and Estella's methodology produces slightly different ending dates for the 1988–89 episode (March 1989); the 1994–95 episode (April 1995); and the 1999–2000 episode (July 2000).

Table 1 shows that the magnitude of tightening actions—as measured by the increase in the nominal FFTR—varied across episodes. The average increase was slightly less than 300 basis points, with a maximum increase of 425 basis points and a minimum increase of 175 basis points. The table also reveals that there were relatively few dissents at the time of liftoff and at the final tightening action.

2.1 Treasury Market Yields During Tightening Episodes

Financial market participants routinely price financial assets like Treasury securities on the basis of current and expected changes in monetary policy. Figure 1 (A–F) shows how short- and long-term Treasury yields changed during each of the six tightening episodes. Each chart plots the daily close of the FOMC's FFTR, the three-month Treasury-bill constant maturity yield, and the 10-year Treasury note constant maturity yield. Vertical lines denote the initial and final actions by the FOMC to increase its FFTR. Implicitly, each chart also reveals the evolution of the term spread, or yield curve.

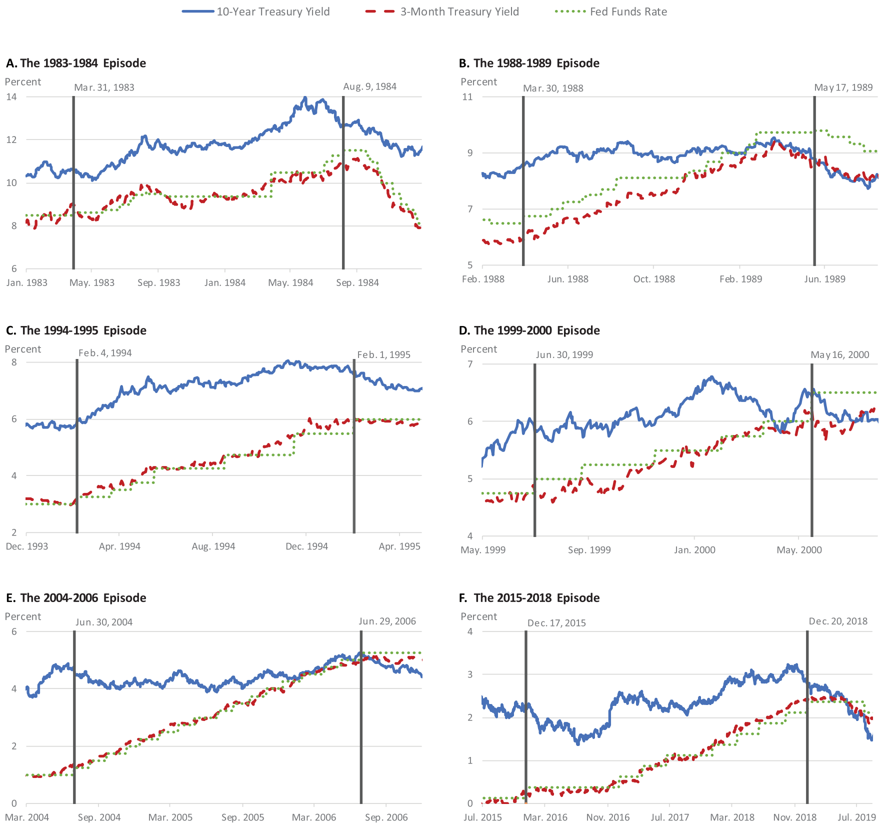
There are several observations to be gleaned from the charts in Figure 1. First, from a broad perspective, nominal short- and

Table 1. FOMC Tightening Actions and Dissents, 1983 to 2018

First Tightening Action	Initial FFR Target (%)	Final Tightening Action	Final FFR Target (%)	Total Tightening (Percentage Points)	Dissent at Initial Action?	Dissent at Final Action?
March 31, 1983	8.5	August 9, 1984	11.5	3.00	None	Yes (1) Wallich wanted easier policy
March 29, 1988	6.50	May 16, 1989	9.8125	3.31	Yes (1) Seger wanted easier policy	Yes (1) Melzer wanted easier policy
February 4, 1994	3.00	February 1, 1995	6.00	3.00	None	None
June 30, 1999	4.75	May 16, 2000	6.50	1.75	Yes (1) McTeer thought tightening was unnecessary	None
June 30, 2004	1.00	June 29, 2006	5.25	4.25	None	None
December 16, 2015	0 to 0.25	Dec. 19, 2018	2.25 to 2.50	2.25	None	None

Source: Board of Governors of the Federal Reserve System and Federal Reserve Bank of St. Louis.

Figure 1. Changes in Short- and Long-Term Treasury Yields During the Six Tightening Episodes



long-term Treasury interest rates rose during the duration of each tightening episode—that is, from the initial increase to the final increase in the FFTR. However, the increases varied across episodes. Second, the three-month bill rate closely followed the path of the FFTR. However, there were periods in the 1988–89 and 1999–2000 episodes when the three-month rate traded below the FFTR for several months. The opposite pattern held over the last several months of the 2015–18 episode, when the three-month yield traded above the FFTR. Third, the average increase in the 10-year nominal Treasury rate across episodes (95 basis points) was appreciably less than the average increase in the FFTR. However, there was considerable variance—from only 23 basis points in the 1988–89 episode to 199 basis points in the 1983–84 episode.

Two of the tightening episodes triggered very different behavior in the long-term Treasury bond market. The first was the 1994–95 tightening period. This episode is unique in Federal Reserve monetary history for a couple of reasons. First, the economy was strengthening in 1994, as real GDP increased from 2.6 percent in 1993 to 4.1 percent in 1994, but there were few obvious inflation pressures. From 1992 to 1994, CPI inflation slowed from 3.1 percent to 2.6 percent.[6] However, the Board staff forecasted that inflation would remain above 3 percent in 1994 (3.3 percent) and in 1995 (3.1 percent).[7] As a result, the Greenspan FOMC adopted a "more radical approach: moving gently and preemptively, before inflation even appeared."[8] A second aspect of this episode was the form of forward guidance the Committee would use to signal pending policy changes in the FFTR at future meetings. For example, Greenspan signaled the Committee's intent a month before—in congressional testimony—to increase the federal funds target rate. Then, at the conclusion of the February 4, 1994, meeting, the FOMC released for the first time a post-meeting statement.[9] The FOMC continued to issue post-meeting statements over the next few years, but only at meetings where a policy change occurred. Beginning with the May 18, 1999, meeting, statements were released after each FOMC meeting.[10] Beginning with the June 30, 1999, statement, the FOMC began to report the target for the federal funds rate.

[6]Unless noted otherwise, yearly changes in output and inflation are reported as percent changes from the fourth quarter of one year to the fourth quarter of the following year.

[7]Reported in the January 28, 1994, Greenbook, Part 1, pp. 1–12.

[8]Greenspan (2007, p. 154).

[9]Despite these signals, financial markets appeared to be taken by surprise, as the subsequent turmoil was termed "the bond market massacre" by *Fortune* magazine. See http://fortune.com/2013/02/03/the-great-bond-massacre-fortune-1994/. By contrast, Borio and McCauley (1995) examined bond market volatility across several countries and found little evidence that the volatility stemmed from actions by monetary or fiscal policymakers.

[10]At the May 18 meeting, Greenspan proposed including a "tilt" in the statement, which, in his view, allowed the FOMC "to move in light of a lot of small indications in the CPI that may suggest a rise in inflation." See the May 18, 1999, FOMC Transcript, p. 58. This tilt, whether intentional or not, signaled the start of the 1999–2000 tightening episode at the following meeting in June.

A markedly different set of circumstances in the Treasury market occurred during the fifth tightening episode. This episode, which began in June 2004 and ended in June 2006, was dubbed the "lower for longer" period. Over this two-year period, encompassing 17 meetings, the FOMC raised its federal funds target rate from 1 percent to 5.25 percent in increments of 25 basis points. However, from May 2004 to early January 2006, long-term Treasury yields remained within a fairly narrow trading range—roughly between 4 percent and 5 percent—despite the steady increase in the FFTR.[11] This event was subsequently termed "the Conundrum" by former Fed Chairman Alan Greenspan.[12] Eventually, long-term yields turned sharply higher, rising by a little less than 100 basis points from mid-January 2006 to early July 2006.

A fourth observation is a common—though not uniform—pattern in the Treasury market during the duration of a Fed tightening action. Namely, short-term rates eventually rise by more than long-term rates, leading to, first, a gradual flattening of the yield curve and then, second, an inversion of the yield curve. Indeed, yield curve inversions occurred in three of the six episodes 1988–89, 1999–2000, and 2004–06. In each of these three episodes, the FOMC increased its FFTR *after* the yield curve inversion. The Committee's behavior was consistent with the asymmetric objective function highlighted in the Greenspan quote above.

A fifth observation from the charts in Figure 1 pertains to the final tightening episode. Prior to the final tightening action on December 20, 2018, 10-year Treasury yields were falling, resulting in a flattening of the yield curve. But unlike three of the previous five episodes, the FOMC's final tightening move occurred before the inversion of the yield curve. Indeed, there was much commentary among FOMC participants about the causes and significance of the flattening yield curve during 2018.[13] One might conjecture that the Committee, recalling the earlier episodes when the FOMC continued

[11] There was some parallel with the 1988–89 episode. Long-term rates peaked early in the tightening cycle, at about 9.25 percent during the week ending May 27, 1988. The 10-year yield would not surpass this level until the week ending August 12; however, over this period the FOMC would raise the federal funds target rate by 113 basis points.

[12] See Thornton (2018) for discussion and analysis of this event.

[13] See, for example, the Minutes of the July 31–August 1, 2018, FOMC meeting.

to raise its FFTR after an inversion of the yield curve, refrained from taking similar action. Regardless, the yield curve would eventually invert in 2019, but well after the final tightening action.[14]

3. Evolving Economic Conditions and Forecast Accuracy during Tightening Episodes

Policymakers confronted a changing macroeconomic environment over the periods encompassed by these tightening episodes that bore little resemblance to episodes before 1983. For example, from 1983 to the start of the pandemic in early 2020, the FOMC was routinely confronted by lower inflation and lower unemployment rates during expansionary periods compared with most episodes that occurred during the 1960s and 1970s. The former period has come to be known as the Great Moderation, while the latter period is known as the Great Inflation (high and rising unemployment rates). Figure 2 captures another key element of the post-1983 economic environment: The steady decline in the natural (real) rate of interest (r*).[15] In monetary policymaking, r* is often used as a benchmark for measuring the stance of policy.[16] All else equal, lower inflation and a lower r* meant that the peak nominal FFTR was sequentially lower during each tightening episode, as seen in Figure 1A–F.

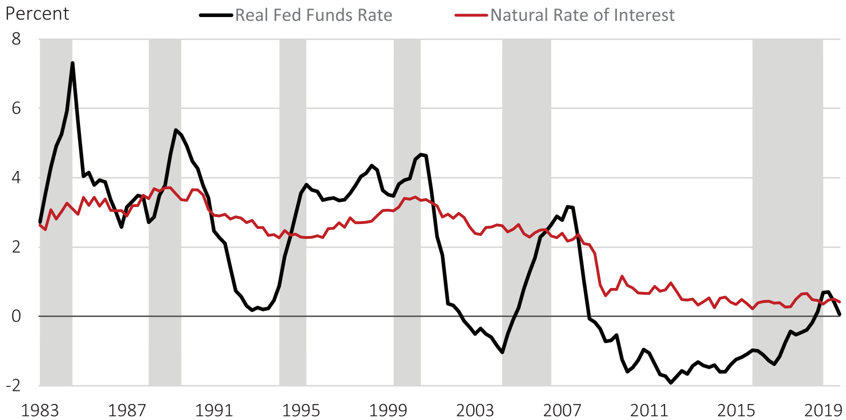
Figure 2 shows that in the first part of the 1991–2001 expansion, the 2001–07 expansion, and the 2009 to 2020 expansion, the real FFTR was well below r* for extended periods. By contrast, the

[14]The yield curve inverted briefly in late March 2019. It would remain inverted from May 23, 2019, to October 10, 2019.

[15]The natural rate of interest (r-star, or r*) is calculated by Holston, Laubach, and Williams (2017). In monetary policymaking, r* is a time-varying, model-based estimate of the real short-term interest rate required to keep inflation and the unemployment rate at the FOMC's target rates. Estimates of r* are imprecise (i.e., have wide confidence bands). Holston, Laubach, and Williams (2017) estimate the sample average standard error is 1.1 percentage points. Their published point estimate for r* in 2016 was 0.4 percent. Holston, Laubach, and Williams suspended the reporting of their r* measure following the onset of the pandemic in early 2020.

[16]If the real FFTR is below r*, then policy is deemed accommodative, leading to faster output growth, falling unemployment rates, and rising inflation pressures. The opposite would occur if the real FFTR was above r*. The real federal funds target rate is defined as the nominal rate less the four-quarter percent change in the PCE price index excluding food and energy.

Figure 2. The FOMC's Real Federal Funds Target Rate and an Estimate of the Natural Rate of Interest (r*) During Tightening Episodes

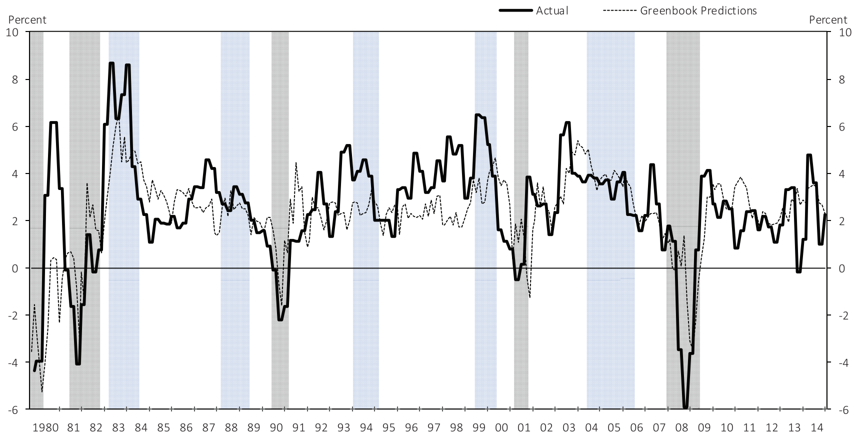


Source: Federal Reserve Bank of St. Louis and Haver Analytics.
Note: Gray shading indicates periods of Fed tightening actions.

real FFTR and r* were more closely aligned before the onset of the 1988–89 tightening, and the real FFTR was well above r* for about seven years from mid-1994 to early 2001; the latter episode is discussed in more detail below. However, in all cases the tightening actions resulted in an eventual modest overshooting of the policy rate (FFTR > r*), leading to a recession in five of the six tightening episodes (the exception being the 1994 episode). One could argue that the recession following the sixth episode was not triggered by tighter monetary policy, but by sharply higher oil prices and the collapse in aggregate economic activity associated with the COVID-19 pandemic.[17] Still, the real FFTR eventually exceeded r*, but by much less than in the previous tightening episodes.

[17]Many private professional forecasters prior to the pandemic had noted the elevated probability of a recession in 2020 according to the December 2019 Blue Chip Consensus forecast. According to the forecast consensus, the probability of a recession in 2020 was about 33 percent. This probability was marginally lower compared with a year earlier when the consensus placed a roughly 35 percent probability of a recession in 2020.

Figure 3. Actual and Predicted Real GDP Growth



Source: Hetzel (2008).
Note: Observations correspond to FOMC meetings. Predictions are from the Greenbook and are for the annualized two-quarter rate of growth of real output (GNP before December 1991 and GDP thereafter). If an FOMC meeting is in the first two months of a quarter, the predicted growth rate is for the contemporaneous and succeeding quarter. If it is in the last month of a quarter, the predicted growth rate is for the succeeding two quarters. Actual growth is the subsequently realized growth rate, measure using the data available at the time of the publication of the "final" GDP estimate for the final quarter of the two-month growth rate. The final estimate is released in the last month of the quarter following a particular quarter. Blue shading indicates periods of Fed tightening. Gray shading indicates a recession.

Trends in the macroeconomy shown in Figure 2 are hard to spot on a meeting-by-meeting basis. Indeed, FOMC policymakers are regularly challenged because shocks and other factors that might alter the structural trends in the macroeconomy are difficult to identify in real time. We can see this in Figure 3, which shows the evolution of actual real GDP growth and the Board staff's projection of real GDP growth during the periods before, during, and after the first five tightening episodes (blue-shaded intervals). The growth rates plotted are annualized two-quarter percent changes. The Board staff's forecasts are reported in the Greenbook and the horizon extends well beyond two quarters.[18] I use two-quarter-ahead forecasts for two

[18]Greenbook forecasts—today known as Tealbook A forecasts—inform FOMC policymakers about the staff's expected short-to-medium-term path of key

primary reasons. First, the data set was readily available. Second, a reading of many FOMC transcripts during the Great Moderation period suggests that monetary policymakers tended to tether their policy discussions and decisions to economic developments that have transpired over the intermeeting period and their implication for the economy over the next six to nine months. Third, the forecasting literature suggests considerable erosion in forecast accuracy as the forecast horizon lengthens. For example, Breitung and Knüppel (2021) use Diebold-Mariano-type and encompassing tests to examine six-quarter-ahead macroeconomic forecasts from Consensus Economics. They find that the information content for quarter-to-quarter forecasted changes in U.S. real GDP growth reaches a maximum at two to three quarters ahead.[19]

Some notable patterns are apparent in Figure 3. First, in the first four tightening episodes (blue-shaded intervals), the Greenbook forecasts generally underestimated the strength of real GDP growth. However, this pattern did not prevail in the fourth episode. In terms of the post-tightening period, which were recessionary periods (gray-shaded intervals) in all except the 1993–94 tightening episode, there generally does not appear to be a consistent pattern. Perhaps not surprisingly, the Greenbook forecasts did not foresee the timing and depth of the Great Recession and Financial Crisis (the period following the fourth tightening episode). Belongia and Ireland (2018) examined Greenbook forecasts from 1987 to 2012 and argued that the FOMC set the FFTR in a manner consistent with the Board staff's forecast for the output gap and inflation. They further argued that the FOMC was less responsive to Greenbook forecasts around turning points in the business cycle.

Visual inspection of actual and projected outcomes can be informative but may mask the true accuracy of the forecasts that

economic indicators, including the nominal and real FFTR. The author thanks Robert Hetzel for sharing the data plotted in Figure 3. The sixth tightening episode is not shown because Tealbook forecasts are released with a five-year lag.

[19]See also Reifschneider and Tulip (2019), who compare the predictive accuracy of Federal Reserve Board staff with FOMC participants and other forecasters (e.g., Congressional Budget Office and Blue Chip) from 1996 to 2015. They find that uncertainty about the economic outlook is "quite large" and that the predictive differences across forecasters for key economic variables like real GDP growth and inflation are "quite small."

Table 2. Greenbook Forecast Accuracy Before, During, and After Tightening Episodes (root mean square errors, %)

	One Year Before Tightening Period	During Tightening Period	One Year After Tightening Period
Medians for Five Episodes	2.6	1.6	1.2
Maximum RMSE (Episode)	2.7 (1983–84)	2.9 (1983–84)	2.3 (1999–2000)
Minimum RMSE (Episode)	1.7 (2004–06)	0.6 (1988–89)	0.8 (1988–89)

Source: Author's calculations based on data plotted in Figure 4.

policymakers relied upon. Table 2 measures Greenbook forecast accuracy (actual less predicted) over three intervals for each of the tightening episodes from the projections and actual outcomes shown in Figure 3. Specifically, the table shows the root mean-squared forecast error (RMSE) for (i) four quarters preceding the beginning of the tightening episode; (ii) the period during the tightening episode; and (iii) the year following the tightening episode. The medians of the five episodes are reported, along with the maximum and minimum RMSE.

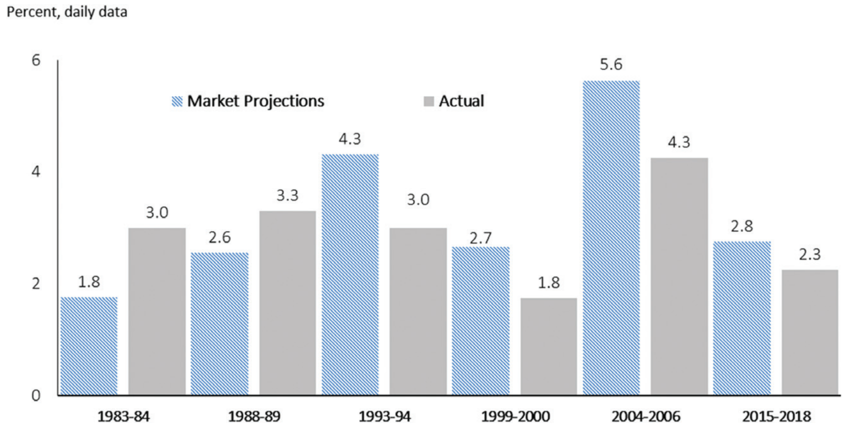
Table 2 shows that Greenbook forecasts for real GDP growth were least accurate (highest RMSE) during the four quarters prior to the beginning of the tightening policy. The median RMSE for the five episodes was 2.6 percent, which was about 63 percent larger than the RMSE during the tightening period and more than twice as large as the median RMSE during the one-year period following the end of the tightening episode. The bottom part of Table 2 shows that the Greenbook's forecast accuracy varied. However, the largest RMSEs were generally associated with the 1983–84 episode, while the smallest RMSEs were generally associated with the 1988–89 episode.

4. Assessing Financial Market Expectations during Tightening Episodes

One of the challenges in measuring responses of firms, households, and financial market participants to the Fed's tightening actions is accounting for expectations about the timing and magnitude of these actions. Regarding the timing of liftoff, the FOMC over time has sought to minimize disruptions to financial markets and economic activity by improving the communication of pending actions to change policy (or not change policy). The 2013 taper tantrum episode is viewed as a counterexample. Improved communication has sometimes taken the form of forward guidance about the future path of the FFTR. If successful, forward guidance can help bring private-sector expectations into closer alignment with the FOMC's intentions, thereby enhancing the effectiveness of monetary policy. Consistent with this view, Poole (2005) found that policy decisions by the FOMC since 1994 elicited little news in the federal funds futures markets. This finding suggests that markets had successfully priced in pending policy decisions by the FOMC. Poole (2005) also found evidence that market expectations of future Fed policy actions were informed importantly by news in the monthly employment report following the introduction of "forward-looking" language in the August 2003 FOMC statement. In a similar vein, Swanson (2006) shows that forecasts of the FFTR by financial markets and private-sector forecasters have become more accurate since the Federal Reserve began a concerted effort since the late 1980s to improve the quantity and quality of its public communication (transparency).

Beyond the scope of the Fed's actions on the expectations of financial markets and forecasters, there is also the issue of how much the Fed should tighten. The magnitude of the Fed's tightening action depends on several factors. This includes, most obviously, the evolving state of the U.S. economy. Markets could be assessing the future state of the economy, and then mapping this trajectory into a well-known policy rule like the Taylor rule. But markets also condition their bets on future policy actions by the FOMC. Markets might also employ a rule of thumb or a heuristic based on previous tightening episodes, or communication from FOMC policymakers about the future stance of monetary policy. These bets are then priced into financial market prices. Gürkaynak, Sack, and Swanson (2005) use high-frequency (intraday) data to show that monetary policy announcements explain a very large variation in long-term Treasury yields that work through the expectations of future policy actions that are reflected in federal funds futures and eurodollar futures rates one-year out.

Figure 4. Magnitude of Tightening Episodes: Actual vs. Expectations



Source: CBOT, CME, FRB.

Figure 4 provides some assessment of the market's expectations about the magnitude of the FOMC's tightening actions relative to the actual amount of Fed tightening. In this case, market expectations are measured at the beginning of the tightening episode using expected future yields derived from three-month eurodollar contracts. No attempt is made to adjust expectations *during* the tightening episode. Specifically, the market's projected tightening in Figure 4 is the terminal value of the farthest traded three-month eurodollar contract less the initial federal funds rate immediately prior to liftoff. Moreover, these eurodollar rates are adjusted for risk premium effects.[20]

Figure 4 shows that markets underestimated the magnitude of the tightening in the first two tightening episodes by about 125 and 75 basis points, respectively. However, in the final four episodes,

[20]Yields on eurodollar futures are adjusted by subtracting the estimated forward-swap rate for a given period and time to maturity. The adjustment averages 25 basis points in the 1988, 1994, and 1999 episodes, 15 basis points in the 2004 episode, and 24 basis points in the 2015 episode.

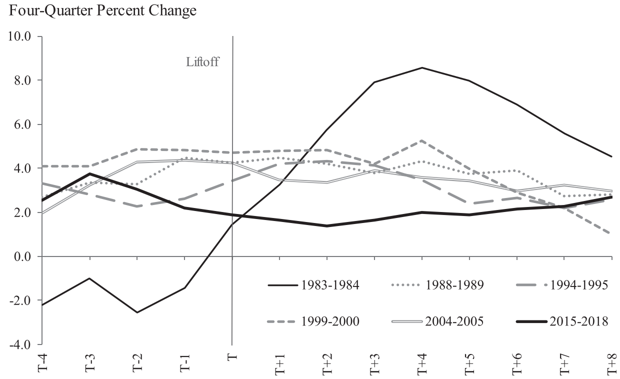
financial market participants overestimated the amount of tightening. The overestimates were especially pronounced in the 1993–94 and 2004–06 episodes—about 130 and 140 basis points, respectively. As noted above, the earlier episode was unique because the Fed's preemptive approach appeared to catch the market by surprise, while the latter episode was unique because of its duration. It is important to emphasize, though, that these expectations were conditioned on the current and prospective state of the economy and other factors that prevailed at the time of liftoff. Finally, it was also the case that markets overestimated the total tightening in the 2015–18 episode, but by much less than the previous three episodes.[21] This is consistent with the literature cited above that the FOMC's shift to a more transparent communication paradigm provided markets with better information than in previous episodes.

5. Key Economic Indicators Before and After Past Tightening Episodes

The remainder of this article will examine the behavior of six key economic variables before and after the onset of tighter monetary policy in the six episodes highlighted in this paper. The six economic indicators are those that are generally of most interest to monetary policymakers: (i) the four-quarter growth in real GDP; (ii) the level of the unemployment rate; (iii) the four-quarter growth of the personal consumption expenditures price index, or PCEPI; (iv) the four-quarter growth of the personal consumption expenditures price index excluding food and energy prices (core PCEPI). These four series are reported in the FOMC's quarterly Summary of Economic Projections. The fifth and six series are measures of short- and long-run inflation expectations: (v) the University of Michigan survey of household inflation expectations over the next 12 months (median estimate); and (vi) the 10-year-ahead forecast for PCEPI

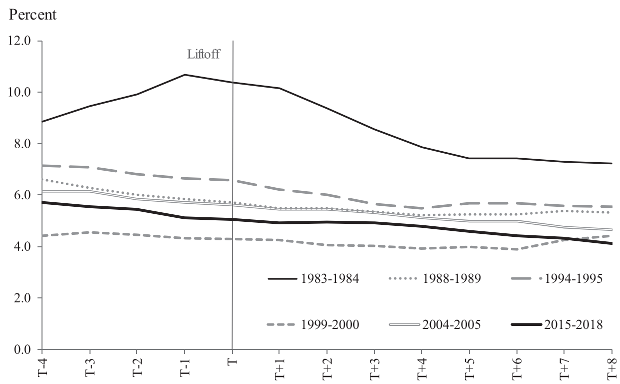
[21] At the December 2015 FOMC meeting, the median participant projected that the federal funds rate would increase from 0.4 percent at the end of 2015 to 3.3 percent in 2018. This cumulative projected tightening (2.9 percentage points) was very close to the expectations of financial market participants (2.8 percentage points).

Figure 5. Real GDP



Source: BEA/Haver Analytics.

Figure 6. Unemployment Rate



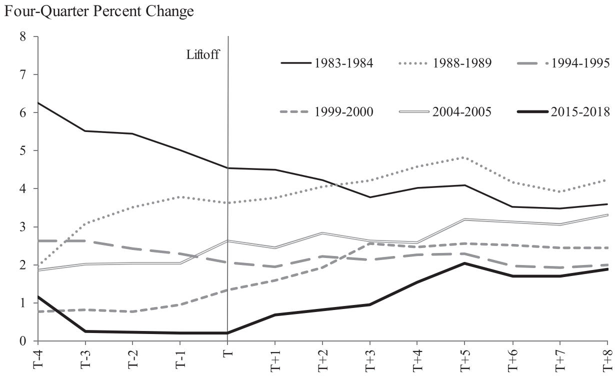
Source: BEA/Haver Analytics.

inflation based on the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters and other sources.[22]

The analysis in this section is based on Figures 5–10 and it will largely be descriptive in nature. The visual representation of the

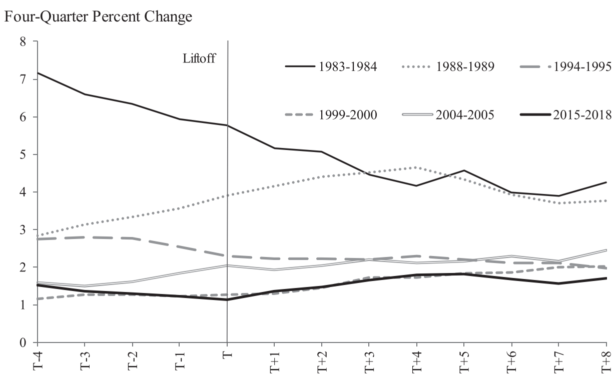
[22]The data used in Figures 5–10 and throughout the article, unless indicated, are current-vintage data—that is, data vintages that existed at the end of 2021.

Figure 7. Personal Consumption Expenditures Price Inflation



Source: BEA/Haver Analytics.

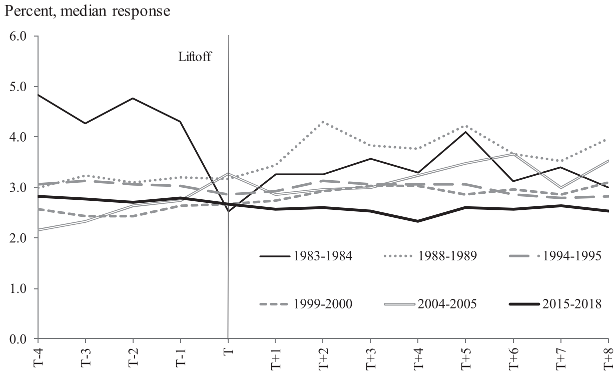
Figure 8. Core Personal Consumption Expenditures Price Inflation



Source: BEA/Haver Analytics.

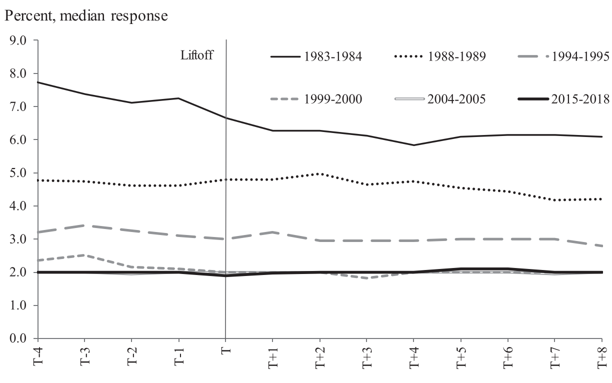
data in Figures 5–10 is a common method of comparing economic activity for some period before and after an arbitrary dividing point. For this paper, the vertical dividing point is the quarter when the FOMC began its tightening episode by raising its FFTR. The figures

Figure 9. 12-Month-Ahead Inflation Expectations (consumers)



Source: University of Michigan/Haver Analytics.

Figure 10. 10-Year-Ahead PCE Inflation Expectations (forecasters)



Source: Board of Governors of the Federal Reserve System/Federal Reserve Bank of Philadelphia/Haver Analytics.

show values four quarters before liftoff and eight quarters after liftoff for each of the six tightening episodes (six lines in each chart).

The economic conditions the prevailed before and after the first and sixth tightening episodes were distinctly different than the other four episodes.[23] In the first episode (1983–84), the four-quarter growth of real GDP (Figure 5) was modestly negative in the three quarters prior to liftoff, and the unemployment rate (Figure 6) had risen, on net, to a peak of about 10.75 percent one quarter prior to liftoff. However, headline (Figure 7) and core (Figure 8) inflation and short- and long-term inflation expectations (Figures 9 and 10) were declining rapidly. In the first episode, actual inflation and long-term inflation expectations continued to decelerate about a year after liftoff; however, short-term inflation expectations rebounded after liftoff. As the Record of Policy Actions for the March 29, 1983, FOMC meeting in Table 3 details, the participants believed that the recovery was under way, though with appreciable uncertainty. As it turned out, the economy continued to rebound strongly after liftoff, with real GDP growth averaging about 6.3 percent in the subsequent two-year period. The Committee also generally thought that the recent rapid growth in the monetary aggregates did not have a material effect on the outlook for inflation.

The sixth episode was similarly unique. Prior to the liftoff in December 2015, real GDP growth had slowed from about 4 percent in early 2015 to about 2 percent during the liftoff quarter. However, the unemployment continued to drift lower in the four quarters prior to the liftoff quarter. Interestingly, headline PCEPI inflation was anchored close to zero before liftoff, while core PCEPI inflation was slowing from about 1.5 percent to 1 percent. Short-term inflation expectations were also drifting lower before liftoff, while long-term inflation expectations remained anchored at 2 percent. Importantly, the sharp slowing in inflation and inflation expectations in 2015 reflected a 45 percent decline in nominal crude oil prices between the fourth quarter of 2014 and the fourth quarter of 2015. Thus, at the time of the liftoff, both headline and core inflation were well below the FOMC's 2 percent target rate. In the Summary of Economic Projections released at the conclusion of the December 2015 meeting, the median FOMC participant projected that core inflation would remain under 2 percent through the end of 2017. Thus,

[23]The working paper version of this article details the responses by several other economic and financial market indicators not detailed here (i.e., the major components of PCE, business loans, equity prices, and business loans). See https://research.stlouisfed.org/wp/more/2020-003.

Table 3. FOMC Description of Policy Decision at First Tightening Action

First Tightening Action	FOMC Description of Policy Action
March 31, 1983	In the Committee's discussion of the economic situation and outlook, the members agreed that a recovery in economic activity appeared to be under way, although several commented that the evidence available thus far was too fragmentary to permit a firm evaluation of the strength of the upturn. While the staff projection of moderate growth for 1983 as a whole was cited as a reasonable expectation, members commented on the many uncertainties surrounding the economic outlook and expressed differing views regarding the direction of possible deviations from the staff projection. . . In discussing a policy course for the weeks immediately ahead, Committee members recognized that substantial uncertainties affected both the economic outlook and the interpretation of the monetary aggregates. Concern was expressed about the implications of the rapid growth in the monetary aggregates, particularly if it should continue. However, it was also noted that the rapid expansion of recent months, given the distortions related to various institutional changes, probably did not have the significance for future economic and price developments that it might have had in the past. (Record of Policy Action, March 29, 1983)
March 29, 1988	In the Committee's discussion of the economic situation and outlook, the members generally agreed that the information available since the February meeting pointed to a stronger expansion in business activity than they had anticipated earlier. Unfortunately, recent developments in the view of several members also increased the risks of more pressures on productive resources and more inflation. A number of members noted that the revised staff forecast was in line with their own projections and some also indicated that any deviations were likely to be in the direction of somewhat faster expansion and even higher rates of inflation. (Record of Policy Action, May 20, 1988)
February 4, 1994	In this situation, the members agreed that monetary policy should be adjusted toward a more neutral stance that would encourage sustained economic growth without a buildup of inflationary imbalances. The members recognized that timely action was needed to preclude the necessity for more vigorous and disruptive policy moves later if inflationary pressures were allowed to intensify. The history of past cyclical upswings had demonstrated the inflationary consequences and adverse effects on economic activity of delayed anti-inflation policy actions. (FOMC Minutes, Released March 25, 1994)

(*continued*)

Table 3. (Continued)

First Tightening Action	FOMC Description of Policy Action
June 30, 1999	Last fall the Committee reduced interest rates to counter a significant seizing-up of financial markets in the United States. Since then much of the financial strain has eased, foreign economies have firmed, and economic activity in the United States has moved forward at a brisk pace. Accordingly, the full degree of adjustment is judged no longer necessary. Labor markets have continued to tighten over recent quarters, but strengthening productivity growth has contained inflationary pressures. Owing to the uncertain resolution of the balance of conflicting forces in the economy going forward, the FOMC has chosen to adopt a directive that includes no predilection about near-term policy action. The Committee, nonetheless, recognizes that in the current dynamic environment it must be especially alert to the emergence, or potential emergence, of inflationary forces that could undermine economic growth. (FOMC Meeting Statement)
June 30, 2004	The Committee believes that, even after this action, the stance of monetary policy remains accommodative and, coupled with robust underlying growth in productivity, is providing ongoing support to economic activity. The evidence accumulated over the intermeeting period indicates that output is continuing to expand at a solid pace and labor market conditions have improved. Although incoming inflation data are somewhat elevated, a portion of the increase in recent months appears to have been due to transitory factors. The Committee perceives the upside and downside risks to the attainment of both sustainable growth and price stability for the next few quarters are roughly equal. With underlying inflation still expected to be relatively low, the Committee believes that policy accommodation can be removed at a pace that is likely to be measured. (FOMC Meeting Statement)
December 15, 2015	Information received since the Federal Open Market Committee met in October suggests that economic activity has been expanding at a moderate pace. Household spending and business fixed investment have been increasing at solid rates in recent months, and the housing sector has improved further; however, net exports have been soft. A range of recent labor market indicators, including ongoing job gains and declining unemployment, shows further improvement and confirms that underutilization of labor resources has diminished appreciably since early this year. Inflation had continued to run below the Committee's 2 percent longer-run objective, partly reflecting declines in energy prices and in prices of non-energy imports. Market-based measures of inflation compensation remain low; some survey-based measures of longer-term inflation expectations have edged down… The Committee judges that there has been considerable improvement in labor market conditions this year, and it is reasonably confident that inflation will rise, over the medium term, to its 2 percent objective. (FOMC Meeting Statement)

Source: Board of Governors of the Federal Reserve System.

it does not appear that the fear of above-target inflation was at the forefront of the Committee's concerns.

Following liftoff in December 2015, real GDP growth continued to decelerate for two quarters, but then the economy began to pick up speed. The unemployment rate continued to drift lower, but inflation—particularly, headline inflation—began to accelerate, reaching the 2 percent inflation target five quarters after liftoff, and sooner than the Committee expected at the time of liftoff. Still, short-term inflation expectations continued to drift lower a year after liftoff, before rebounding slightly. Table 3 shows that the Committee at the time of liftoff acknowledged the below-target inflation rates and low inflation expectations but was confident that further improvement in labor markets would begin to push inflation higher.

The remaining four tightening episodes were broadly similar—namely, the FOMC was tightening into a strengthening economy, with a falling unemployment rate, and with rising inflation. As noted in Table 3, the FOMC generally noted these developments and were worried about the potential for rising inflation. However, in the 2004 episode, they initially noted that underlying inflation at the time of liftoff was "expected to be relatively low."

5.1 Notable Differences Across Episodes

There were some notable differences across episodes. First, the conventional view that tighter monetary policy eventually leads to slower real GDP growth over the medium term generally only held in the four episodes that spanned 1988 to 2005. In these four episodes, real GDP growth was accelerating modestly in the four quarters before liftoff, remained roughly constant over the four quarters following liftoff, then output growth decelerated, on average, a little more than 1 percentage point in quarters 5–8 following liftoff. Second, the sixth episode was unique in that there was a year-long pause between liftoff and the second increase in the FFTR. In the fourth quarter of 2015, real GDP was up 1.9 percent from four quarters earlier, about a percentage point below the average of the previous four quarters (see Figure 5). Output growth would continue to decelerate modestly four quarters after liftoff, averaging 1.7 percent. But as the economy began to improve in the second half of 2016, and short-term inflation expectations stabilized, the FOMC resumed its tightening actions in December 2016.

A third key difference is the behavior of short- and long-term inflation expectations during the 2015–18 episode. In the first five episodes, on average, short-term inflation expectations accelerated modestly following liftoff. This was consistent with the quote from Greenspan noted earlier, who indicated that the Fed appeared to have an asymmetric objective function—worrying more about a resurgence of inflation from tightening too little, and less about the risk of weaker output growth and employment from tightening too much. And this pattern generally held for long-term inflation expectations as well following liftoff in the first five episodes, as long-term inflation expectations continued to decelerate modestly after liftoff. But this pattern did not hold in the sixth episode.

At the October 2015 meeting, the FOMC concluded that it was appropriate to "wait for additional information" before beginning the normalization process, but also noted that "even after employment and inflation are near mandate-consistent levels, economic conditions may, for some time, warrant keeping the target federal funds rate below levels the Committee views as normal in the longer run." As suggested by the FOMC statement language in Table 3, many FOMC members appeared to be worried about an acceleration in core inflation in the midst of improving labor market conditions (a falling unemployment rate).[24] This suggests that many FOMC members still relied on the Phillips-curve framework to forecast inflation, despite evidence that its usefulness as a guide to policymaking was much less appropriate in the final episode than it was in previous episodes.[25]

[24]The minutes of the December 13–14, 2016, FOMC meeting noted the following: "Many participants judged that the risk of a sizable undershooting of the longer-run normal unemployment rate had increased somewhat and that the Committee might need to raise the federal funds rate more quickly than currently anticipated to limit the degree of undershooting and stem a potential buildup of inflationary pressures."

[25]In a 2018 speech, Chair Powell presented evidence that the slope of the Phillips curve and the inflation persistence coefficient was much lower in 2015 than in previous episodes. See Powell (2018). Using Powell's framework, the slope coefficient in 1994 was estimated to be −0.53 and the persistence coefficient was 1.03. By 2015, the slope coefficient was −0.07 and the persistence coefficient was 0.45. (Note: Author's calculations are available on request.)

Nevertheless, it is apparent from Figures 7–10 that the FOMC faced a conundrum in the final tightening episode. Actual headline and core inflation was rising modestly following liftoff in December 2015, but inflation was still below the 2 percent inflation target. Moreover, short- and long-term inflation expectations were moving modestly lower. It thus appears that the FOMC discounted the signal from inflation expectations and chose instead to rely more on the signal from the modest upswing in actual inflation and, concurrently, that the continued fall in the unemployment rate—via Phillips curve effects—would trigger an acceleration in core inflation. Alas, this development failed to materialize to the degree many policymakers expected.

From a longer-term perspective, Figures 7–8 and 10 show that actual inflation and long-term inflation expectations have trended lower since 1983. Despite generally lower actual and expected inflation over time, as Figure 4 showed, the cumulative increase in the FFTR has not declined nearly as much. The median total increase in the FFTR during these six episodes was 308 basis points—ranging from a low of 175 basis points in the 1999–2000 episode, to a high of 425 basis points in the 2004–06 episode.

6. Conclusion

The decision to undertake a series of tightening actions presents unique challenges for Fed policymakers. Using a variety of economic metrics, this article has examined the six monetary policy tightening episodes pursued by the FOMC since 1983. In the first five episodes, the data clearly suggest that the FOMC was tightening into a strengthening economy, sometimes with a lag, and with building price pressures. As the FOMC continued to tighten, the yield curve eventually inverted in three of the four episodes and the economy subsequently fell into an economic recession. One exception was the 1994–95 tightening episode. In that episode, neither development occurred. The other exception was the 2015–18 episode. Although the U.S. economy fell into a deep recession in the spring of 2020, the primary cause was the direct and indirect effects of the COVID-19 pandemic. The sixth episode was unique in other ways. Probably the most important difference is that the FOMC began its tightening regime following a notable deceleration in real GDP growth, with headline and core inflation remaining well below the FOMC's 2 percent inflation target, and with short- and long-term inflation expectations trending slightly lower.

References

Adrian, T., and A. Estrella. 2009. "Monetary Tightening Cycles and the Predictability of Economic Activity." Staff Report No. 397, Federal Reserve Bank of New York.

Belongia, M. T., and P. N. Ireland. 2018. "Monetary Policy Lessons from the Greenbook." Working Paper No. 955, Boston College.

Borio, C. E. V., and R. N. McCauley. 1995. "The Anatomy of the Bond Market Turbulence of 1994." Working Paper No. 32, Bank for International Settlements. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=11153.

Breitung, J., and M. Knüppel. 2021. "How Far Can We Forecast? Statistical Tests of the Predictive Content?" *Journal of Applied Econometrics* 36 (4): 369–92.

Greenspan, A. 2007. *The Age of Turbulence: Adventures in a New World*. Penguin Press.

Gürkaynak, R. S., B. Sack, and E. T. Swanson. 2005. "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements." *International Journal of Central Banking* 1 (1, May): 55–94.

Hetzel, R. L. 2008. *The Monetary Policy of the Federal Reserve: A History*. Cambridge University Press.

Holston, K., T. Laubach, and J. C. Williams. 2017. "Measuring the Natural Rate of Interest: International Trends and Determinants." *Journal of International Economics* 108 (Supplement 1): S59–S75.

Laforte, J.-P., and J. Roberts. 2014. "November 2014 Update of the FRB/US Model." FEDS Notes, November 21.

Orphanides, A. 2015. "Fear of Liftoff: Uncertainty, Rules, and Discretion in Monetary Policy Normalization." *Review* (Federal Reserve Bank of St. Louis) 97 (3): 173–96.

Poole, W. 2005. "How Predictable is Fed Policy?" *Review* (Federal Reserve Bank of St. Louis) 87 (6): 659–68.

Powell, J. H. 2018. "Monetary Policy and Risk Management at a Time of Low Inflation and Low Unemployment." Remarks at the 60th Annual Meeting of the National Association for Business Economics, Boston, Massachusetts, October 2.

Reifschneider, D., and P. Tulip. 2019. "Gauging the Uncertainty of the Economic Outlook Using Historical Forecasting Errors: The Federal Reserve's Approach." *International Journal of Forecasting* 35 (4): 1564–82.

Sharpe, S. A., and G. A. Suarez. 2015. "Why Isn't Investment More Sensitive to Interest Rates: Evidence from Surveys." Finance and Economics Discussion Series No. 2004-002, Board of Governors of the Federal Reserve System. https://www.federalreserve.gov/econresdata/feds/2014/files/201402r.pdf.

Swanson, E. T. 2006. "Have Increases in Federal Reserve Transparency Improved Private Sector Interest Rate Forecasts?" *Journal of Money, Credit and Banking* 38 (3): 791–819.

Thornton, D. L. 2018. "Greenspan's Conundrum and the Fed's Ability to Affect Long-Term Yields." *Journal of Money, Credit and Banking* 50 (2–3): 513–43.

Willems, T. 2020. "What Do Monetary Contractions Do? Evidence From Large, Unanticipated Tightenings." *Review of Economic Dynamics* 38 (October): 41–58.

Willis, J. L., and G. Cao. 2015. "Has the U.S. Economy Become Less Interest Rate Sensitive?" *Economic Review* (Federal Reserve Bank of Kansas City) (Second Quarter): 5–36.