Volume 16, Number 2

March 2020

INTERNATIONAL JOURNAL OF CENTRAL BANKING
Volume 16, Number 2
March 2020

IJCB
INTERNATIONAL JOURNAL OF CENTRAL BANKING

International Journal of Central Banking

Volume 16, Number 2 March 2020

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 quarterly (ISSN: 1815-4654). Online access to the publication is available free of charge at **www.ijcb.org**. Changes to hard copy subscriptions and orders should be placed by e-mail at editor@ijcb.org.

Renewals, claims, address changes, and requests for permission to reprint material from this journal should be addressed to:

International Journal of Central Banking
HS08.20
European Central Bank
Sonnemannstr. 22
D-60640 Frankfurt am Main
Germany

Phone: +4969134495894
E-mail: editor@ijcb.org

ISSN: 1815-4654

Introduction

This issue of the *International Journal of Central Banking* includes two of the papers presented at the conference entitled "Commodity Prices and Monetary Policy: New Theory and Evidence" hosted by Norges Bank on June 4–5, 2019. The conference was co-organized with Norges Bank, the Bank of Canada, and the Centre for Applied Macro- and Petroleum Economics. The two papers, chosen using the same rigorous standards applied to all *International Journal of Central Banking* content, are "Oil Prices and Consumption across Countries and U.S. States" by Andrea De Michelis, Thiago Ferreira, and Matteo Iacoviello; and "Monetary Policy, Commodity Prices, and Misdiagnosis Risk" by Andrew J. Filardo, Marco J. Lombardi, Carlos Montoro, and Massimo Minesso Ferrari. The program committee for the conference was Pierpaolo Benigno, Drago Bergholt, Hilde Bjørnland, Oleksiy Kryvtsov, Loretta J. Mester, and Barbara Rossi.

Oil Prices and Consumption across Countries and U.S. States*

Andrea De Michelis, Thiago Ferreira, and Matteo Iacoviello
Federal Reserve Board of Governors

We study the effects of oil prices on consumption across countries and U.S. states, by exploiting the time-series and cross-sectional variation in oil dependency of these economies. We build two large data sets: one with 55 countries over the years 1975–2018, and another with all U.S. states over the period 1989–2018. We then show that oil price declines generate positive effects on consumption in oil-importing economies, while depressing consumption in oil-exporting economies. We also document that oil price increases do more harm than the good afforded by oil price decreases both in the world and in U.S. aggregates.

JEL Codes: Q43, E32, F40.

1. Introduction

The large fluctuations of oil prices in recent years have re-ignited interest regarding the macroeconomic effects of oil price shocks. The conventional wisdom of academics, policymakers, and market practitioners is that declines in oil prices boost global economic activity, as increases in consumption, especially for oil importers, outweigh the negative effects on oil producers. This wisdom has recently been reiterated by Arezki and Blanchard (2014), Bernanke (2016), and even Warren Buffett (CBS/AP 2016), amongst others. However, much of

*The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of anyone else associated with the Federal Reserve System. Erin Markievitz, Lucas Husted, Andrew Kane, Joshua Herman, Patrick Molligo, Brynne Godfrey, Charlotte Singer, and Caitlin Dutta provided outstanding research assistance. We are grateful for the comments by Carola Binder. Author e-mails: andrea.de.michelis@frb.gov; thiago.r.teixeiraferreira@frb.gov; matteo.iacoviello@frb.gov.

the research on the global effects of oil price changes has only looked at a restricted set of countries (see, e.g., Jiménez-Rodriguez and Sánchez 2005), or has focused on a set of macroeconomic variables, such as industrial production or current account, whose responses may not be easily mapped into those of gross domestic product (GDP) and consumption (see, for instance, Aastveit, Bjørnland, and Thorsrud 2015 and Kilian, Rebucci, and Spatafora 2009).

This paper studies the effects of oil prices on consumption and GDP across countries and across U.S. states. A hurdle in obtaining precise estimates of the global effects of oil price shocks is data availability. To address this shortcoming, we put together a quarterly data set containing data on GDP, consumption, and oil dependence for 55 countries. The coverage, which varies across countries, spans more than 40 years from 1975 to 2018. We go to great lengths to ensure that the data are consistent across countries and do not contain breaks. Using this unique database, we estimate the effects of oil price shocks on different countries, allowing for the effects to vary based on each country's oil dependence. We find that oil price declines have large, positive effects on the consumption and GDP of oil-importing countries, while depressing consumption and GDP of oil exporters, and that, in aggregate, the positive effects dominate. However, we also show that the boost to oil importers occurs rather gradually, while the hit to oil exporters is realized fairly quickly, thus suggesting that, in aggregate, the benefits from lower oil prices might occur only slowly. Moreover, we show that oil price decreases generate smaller boosts to the world consumption than the drag from comparable oil price increases.

We conclude by presenting complementary analysis on the heterogeneous effects of oil shocks using state-level data for the United States. To this end, we use state-level data on registrations of new cars—available quarterly from 1989 to 2018—as a proxy for state-level consumption. This data set has important advantages over state-level data on personal consumption expenditures, as the latter are still experimental, are available only annually, and cover a shorter time frame. We then show that the effects of oil price declines on consumption differ across states depending on their oil dependence, despite the fact that these states face common monetary and fiscal policy. The paper proceeds as follows. Section 2 presents the results from an aggregate vector autoregressive (VAR) model which groups countries according to whether they are oil importers or oil exporters. Section 3 uses a cross-country panel VAR to dig deeper into the heterogeneous country responses. Section 4 presents the evidence from a state-level panel VAR for the United States. Section 5 relates the results from this paper to the rest of the existing literature on the macroeconomic effects of oil shocks. Section 6 concludes.

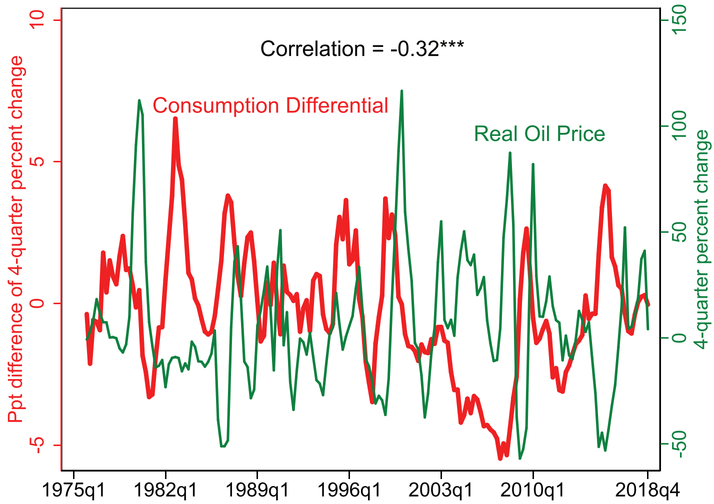
2. Oil Prices, World Consumption, and GDP

Changes in oil prices may reflect disturbances to global aggregate demand as well as disturbances that are specific to the oil market. In most of what follows, we interpret oil price changes as arising from "oil shocks," controlling for shifts in global economic activity that simultaneously drive oil demand and oil prices. We interpret these oil shocks as reflecting disruptions in oil supply due to geopolitical or natural events, or from precautionary or speculative shifts in the demand for oil.

In order to set the stage for the empirical analysis, it is useful to briefly review the key channels by which an "exogenous" decline in oil prices should affect economic activity in both oil-producing and oil-importing countries. On the supply side, lower oil prices raise output in the non–oil sector by reducing firms' production costs and causing investment and output to rise. This cost channel should be stronger in countries or sectors that heavily rely on oil as an input in production. Conversely, falling oil prices may depress energy-related investment across oil-producing countries, dragging down aggregate activity. On the demand side, lower oil prices transfer wealth towards oil-importing countries and away from oil exporters (e.g., Bodenstein, Erceg, and Guerrieri 2011), cause a windfall income gain for consumers, and thus shift consumption towards oil-importing countries. This wealth effect may in turn cause GDP to rise through multiplier effects, and may be larger in sectors that produce goods that are complementary to consumption of oil, such as the automobile sector.[1] Most of our analysis focuses on the consumption effects

[1] See Hamilton (2008) for a discussion on several channels through which oil shocks may affect macroeconomic variables.

Figure 1. Oil Prices and Consumption Differential



Notes: Consumption differential is the aggregate consumption growth of oil importers minus the aggregate consumption growth of oil exporters (black line with y-axis on the left, or red line in online version). Real oil prices (gray line with y-axis on the right, or green line in online version) are measured in 2009 dollars. While consumption differential is measured in percentage-point difference in four-quarter percent changes, real oil prices are measured in four-quarter percent change. The correlation between consumption differential and real oil prices is statistically significant at the 99 percent confidence level (***). Exporters and importers are aggregated using weights based on each country's GDP in constant U.S. dollars. The list of countries and their classification of importer/exporter is in table A.1 in appendix A.

of oil price changes, and exploit the differential exposure of oil producers and oil importers to these changes to highlight its empirical relevance.

Figure 1 shows our first approach to evaluating the differential effects of oil price on consumption according to the country's status as oil importer or exporter. It plots real oil price growth against the "consumption differential," the difference in consumption growth between importing and exporting countries. Since 1975, whenever oil price growth has fallen, the consumption differential has tended to rise, that is, consumption has grown relatively faster in importing countries. The correlation between oil price growth and the consumption differential is −0.32 and is significantly different from zero. It is important to highlight the consumption differential because it offers a simple way to control for global demand-side effects that otherwise create a positive correlation between activity and oil prices. Indeed, separately taken, the correlations of oil prices with importers' and exporters' consumption are both positive, at 0.06 and 0.34, respectively.

2.1 Aggregate VAR: Oil Importers versus Exporters

To further explore the differential responses of oil importers and exporters, we consider a vector autoregressive model aimed at quantifying the effects of shocks to oil prices on global economic activity. The VAR uses two lags and quarterly data from 1975:Q1 to 2018:Q4 on oil prices in real terms (oil_t), global oil production ($oprod_t$), importers' and exporters' private consumption expenditure (ci_t and ce_t, respectively), and importers' and exporters' GDP ($gdpi_t$ and $gdpe_t$, respectively). Consumption and GDP are constructed by aggregating data—using constant-dollar GDP weights—for 45 oil-importing and 10 oil-exporting countries.[2] The countries in our sample, as well as their status as oil importer or exporter, are listed in table A.1 in appendix A.[3] The structural VAR representation takes the form

$$\mathbf{A}\mathbf{x}_t = \mathbf{B}\left(L\right)\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t, \tag{1}$$

$$\mathbf{x}_t = \left[gdpi_t, gdpe_t, ci_t, ce_t, oil_t, oprod_t\right]', \tag{2}$$

$$\mathbb{E}\left(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t{}'\right) = \mathbf{I}, \tag{3}$$

where $\mathbf{B}\left(L\right)$ is a lag polynomial of order 2, and $\mathbb{E}$ is the expectational operator.

We then seek to isolate the effects of oil price fluctuations stemming from oil-market-specific shocks. To do so, we use a Cholesky decomposition of the variance-covariance matrix of the reduced-form VAR residuals whereby oil price shocks affect oil prices and production on impact, and consumption and GDP with a

[2]The sample includes countries which together contribute to about 80 percent of world GDP.

[3]We take the log of all variables and then quadratically detrend them. See appendix A for details on data construction.

one-quarter delay. That is, the reduced-form errors can be decomposed as follows:

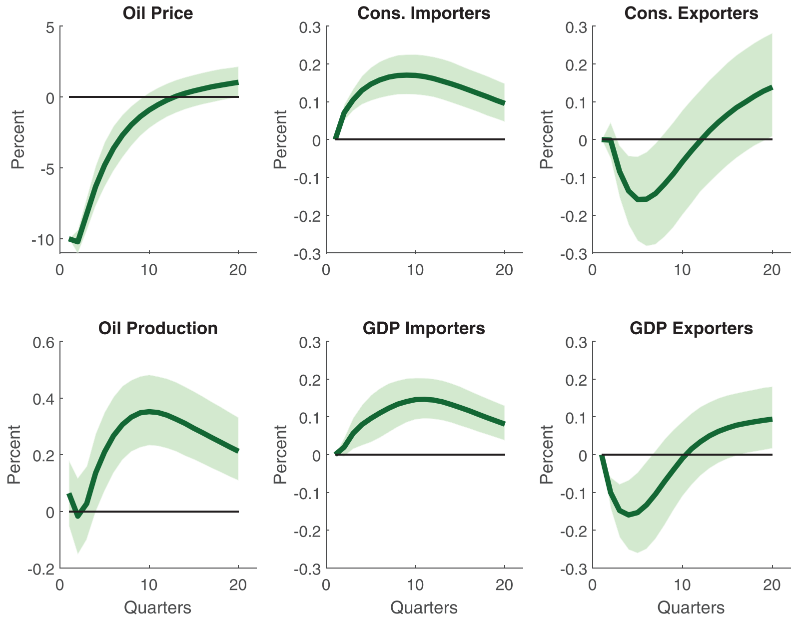
$$\mathbf{e_t} = \mathbf{A}^{-1}\boldsymbol{\varepsilon}_\mathbf{t} = \begin{pmatrix} e_t^{gdpi} \\ e_t^{gdpe} \\ e_t^{ci} \\ e_t^{ce} \\ e_t^{oil} \\ e_t^{prod} \end{pmatrix}$$

$$= \begin{bmatrix} a_{11} & 0 & 0 & 0 & \mathbf{0} & 0 \\ a_{21} & a_{22} & 0 & 0 & \mathbf{0} & 0 \\ a_{31} & a_{32} & a_{33} & 0 & \mathbf{0} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} & \mathbf{0} & 0 \\ a_{51} & a_{52} & a_{53} & a_{54} & \mathbf{a}_{55} & 0 \\ a_{61} & a_{62} & a_{63} & a_{64} & \mathbf{a}_{65} & a_{66} \end{bmatrix} \begin{pmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \\ \varepsilon_t^3 \\ \varepsilon_t^4 \\ \varepsilon_t^{oil} \\ \varepsilon^6 \end{pmatrix}, \qquad (4)$$

where oil price shocks ε_t^{oil} affect the economy according to the fifth column of the $\mathbf{A}^{-1}$ matrix. This assumption is in line with a host of empirical VAR studies that place asset prices after the macroeconomic variables in a Cholesky ordering of the residuals of a VAR, on grounds that asset prices are highly sensitive to contemporaneous economic news or shocks (see, for instance, Bernanke, Boivin, and Eliasz 2005 and Stock and Watson 2005).

Figure 2 shows the impulse response functions (IRFs) to an identified oil price shock that is normalized to imply a 10 percent decline in real oil prices upon the shock impact. This fall in oil prices boosts importing countries' GDP, with a peak effect reached 10 quarters after the shock. Importers' consumption also rises slowly, with a magnitude slightly larger than GDP. The larger increase in consumption relative to GDP suggests that wealth effects may be an important channel by which oil price declines are transmitted to activity in oil-importing countries. By contrast, exporters' consumption drops rapidly and exporters' GDP declines even more so, bottoming out four and three quarters after the oil price shock, respectively. The larger decline in GDP relative to consumption in oil-exporting countries is consistent with a rapid deterioration in investment in the energy sector. Finally, oil production moves little on impact and

Figure 2. Impulse Response Functions to Oil Price Shock in Aggregate VAR



Notes: The figure shows the IRFs to oil price shocks in the aggregate VAR (section 2.1). Solid lines represent median IRFs. Shaded areas represent 16th and 84th percentiles constructed with 5,000 bootstrap replications. Variables are expressed as percent deviation from quadratic trend.

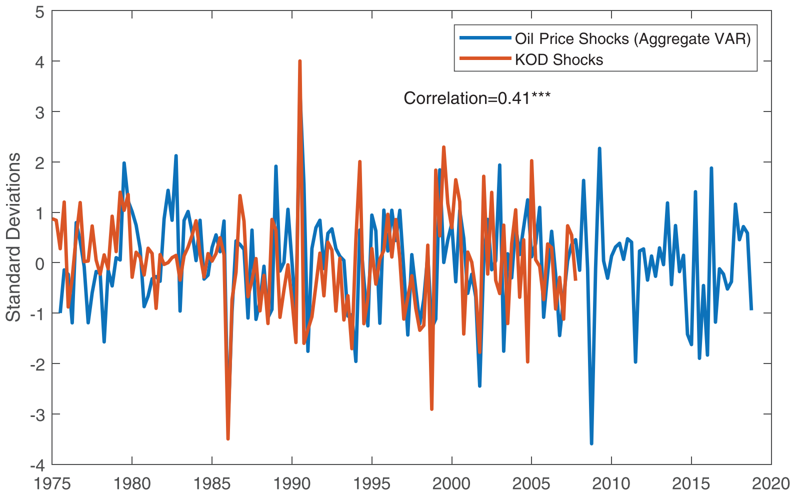
rises persistently in the subsequent quarters, peaking at 0.3 percent above baseline after about two years. Such response supports the interpretation of the oil price shock as being caused by news about higher future supply.

All told, the evidence from the aggregate VAR is consistent with the role of wealth effects on consumption for oil importers, and suggestive of the importance of direct effects on energy-related investment for oil exporters.

2.2 Aggregate VAR under Different Identification Strategies

Our benchmark VAR (equations (1)–(4)) orders oil prices before oil production in the Cholesky decomposition of the variance-covariance matrix of the residuals, with this ordering often being referred to

Figure 3. Comparison of Oil Shocks Identified under Different Strategies



Notes: "Oil Price Shocks (Aggregate VAR)" represent the oil shocks identified by the aggregate VAR from section 2.1 of this paper. "KOD Shocks" represent the "oil-specific demand" shocks identified by Kilian (2009) averaged to the quarterly frequency. The correlation between the two series of shocks is statistically significant at the 99 percent confidence level (***).

as "exogenous oil price assumption" (e.g., Stock and Watson 2016). This assumption regards unexpected changes in oil prices, once world factors are controlled for, as reflecting international developments specific to oil markets, such as exogenous changes in oil supply conditions. A different approach is to order oil production before oil prices, as in Kilian (2009). With this different ordering, it is possible to distinguish oil supply shocks—the residuals of the oil production equation in the VAR—from oil-specific demand shocks—the residuals of the oil price equation in the VAR. In this second ordering, oil-specific demand shocks are a key driver of oil prices and are interpreted as shifts in demand due to concerns about future availability of oil supply.

Figure 3 shows that the two ordering assumptions identify very similar shocks. In particular, the correlation between our identified oil price shocks and Kilian's oil-specific demand shocks is 0.41, and is statistically different from zero. An important reason for the similarity between our estimated shocks and those in Kilian (2009) is that there is little contemporaneous correlation in the data between oil prices and oil production. Therefore, once global conditions are controlled for, it does not seem to matter whether one orders prices before or after production.[4]

Our VAR ordering does not impose any restriction on the contemporaneous co-movement of oil production and oil prices in response to an oil price shock. If one interprets the identified shock as stemming from oil supply, the ratio between the IRF at impact of oil production and the IRF at impact of oil prices measures the implied short-run elasticity of oil demand to oil prices. Given the IRFs from our VAR shown in figure 2, production rises by 0.06 percent when prices drop by 10 percent, thus implying a short-run oil demand elasticity of 0.006, which is in the low end of the estimates of Caldara, Cavallo, and Iacoviello (2019) (henceforth abbreviated to CCI).

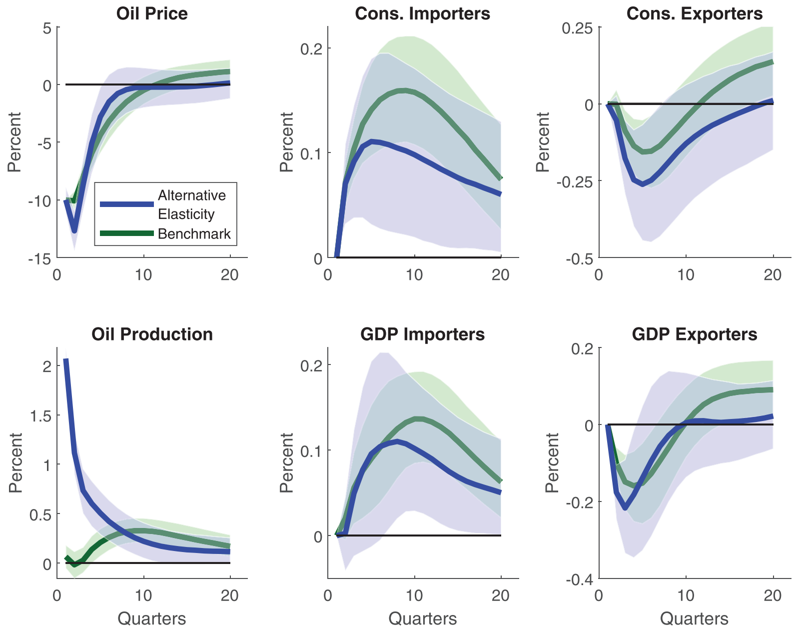
We then examine the robustness of our benchmark results by adopting an alternative identification strategy. We retain the assumption that oil shocks affect GDP and consumption with a one-period delay, but we decompose the variance-covariance matrix of the VAR residuals assuming joint oil demand and oil supply elasticities (the 2×2 block in the bottom right of matrix $\mathbf{A}^{-1}$) in line with existing studies, such as CCI. In particular, we assume an impact oil supply elasticity of 0.1, which in turn implies a VAR-consistent oil demand elasticity of -0.21.[5] As shown in figure 4, when this newly identified oil "supply" shock is scaled so that its effect on oil prices is the same as that of our benchmark specification, the response of oil production is now larger. Nevertheless, the responses of GDP and consumption of oil importers and exporters are similar to those estimated using the benchmark aggregate VAR.

Since the identification strategies from Kilian (2009) and CCI yield results similar to ours, we interpret our results as relatively robust and we maintain a similar order of variables in the VARs

[4]For a detailed discussion, see the appendix in Caldara, Cavallo, and Iacoviello (2019).

[5]Notice that the Kilian (2009) ordering of variables implies a short-run oil supply elasticity of zero, which is not consistent with studies such as CCI.

Figure 4. Impulse Response Functions to Oil Shock under Alternative Oil Supply Elasticity



Notes: "Benchmark" IRFs are calculated using the aggregate VAR of section 2.1 and are shown in green (see online version for colors). "Alternative Elasticity" IRFs, shown in blue, are calculated under the assumption of a world oil demand elasticity of −0.21 and world supply elasticity of 0.1, as explained in section 2.2. Solid lines represent median IRFs. Shaded areas represent 16th and 84th percentiles constructed using 5,000 bootstrap replications. Variables are expressed as percent deviation from quadratic trend.

of the next sections. Importantly, we focus on identifying oil price movements that are driven by shocks that are specific to the oil market, and thus we do not try to separate out fluctuations driven by demand or supply shocks.

3. Oil Dependency across Countries: Panel VAR

In the previous section, we assumed that the effects of an oil price shock only depend on whether a country is an oil importer or exporter, disregarding any cross-sectional and time-series variations in countries' oil dependency. In this section, we account for such variation and show that GDP and consumption of countries with higher oil dependency increase more after exogenous oil price declines. We also show that oil price increases do more harm than the good afforded by oil price decreases.

3.1 Oil Dependency across Countries

Figure 5 documents important differences in oil dependence, both across countries and within country over time. Such differences are obviously lost when we use the simple distinction between oil importers and oil exporters. For a country i at quarter t, we measure oil dependence as the ratio of net oil imports to total oil consumption expressed in percent:

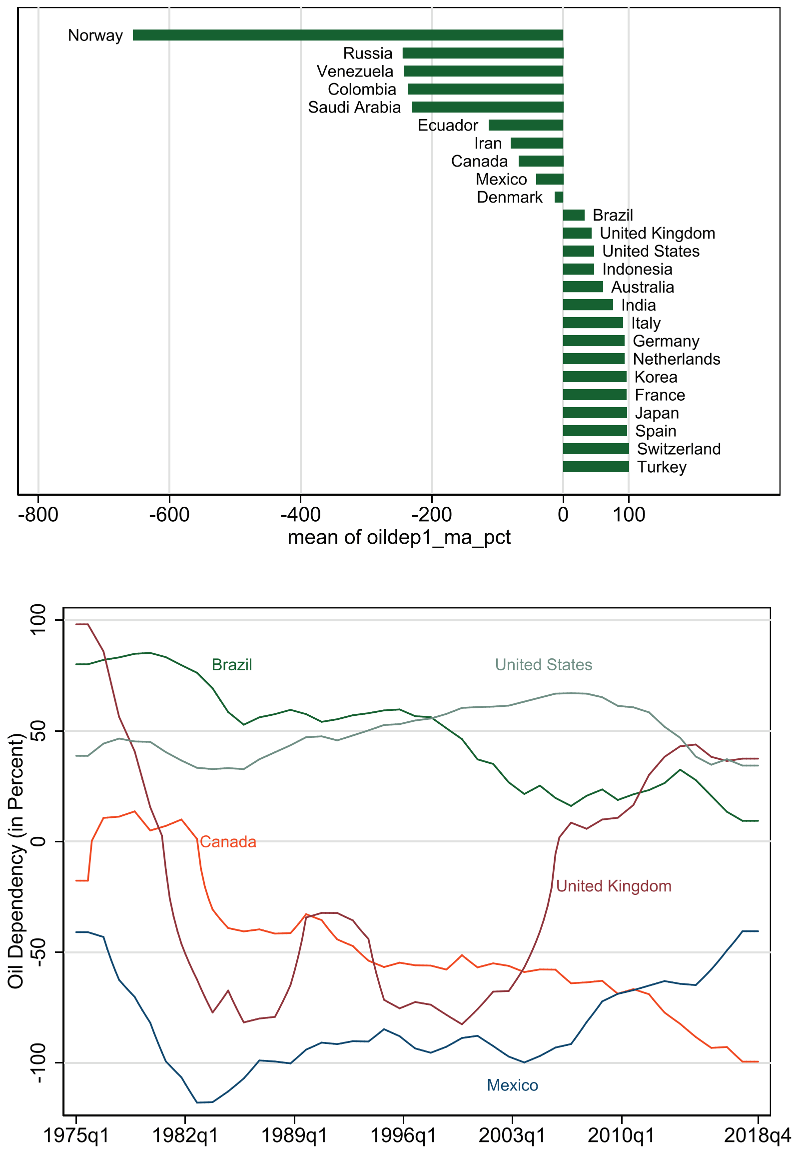
$$d_{i,t} \equiv 100 \times \frac{ocons_{i,t} - oprod_{i,t}}{ocons_{i,t}}, \qquad (5)$$

where $ocons_{i,t}$ is the amount of oil consumed and $oprod_{i,t}$ is the amount of oil produced.

The top panel of figure 5 shows considerable differences in oil dependence across countries in, say, 2013 (similar differences show up when looking at other points in time). Among oil importers, Japan and the large majority of countries in the euro area had an oil dependence of nearly 100 percent, while the United States had an oil dependence of about 50 percent. Among oil exporters, there was even greater heterogeneity according to this metric, ranging from Canada's (net) oil exports of about 70 percent of its oil consumption to Norway's 650 percent.

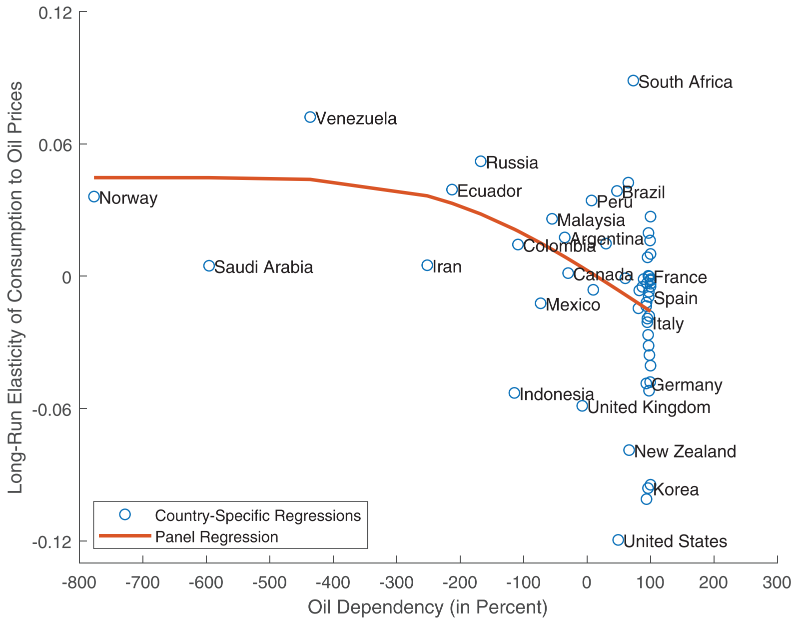
The bottom panel of figure 5 presents evidence on the variation of oil dependency over time for a select group of countries. For instance, Canada was a net importer in the late 1970s and early 1980s, but now exports a large share of its oil production. Additionally, oil dependency in the United States has fallen sharply since 2010, following the massive growth in shale oil production. By contrast, the United Kingdom was a large exporter throughout the 1980s and the 1990s, but is now a net importer. Table A.1 in appendix A provides more details on oil dependency across countries.

Figure 5. Oil Dependency across Countries and over Time



Notes: The top panel shows oil dependency in 2013 for a subset of the largest countries in the sample (largest exporters and countries with share of world GDP in 2013 larger than 1 percent), where oil dependency is measured by oil imports as a percentage of oil consumption (equation (5)). The bottom panel plots oil dependency over time for a select group of economies.

Figure 6. Oil Dependency and Long-Run Elasticity of Consumption to Oil Prices



Notes: On the horizontal axis, oil dependency is calculated as the sample average, where oil dependency is measured by oil imports as a percentage of oil consumption (equation (5)). The vertical axis features long-run elasticities of consumption to changes in oil prices. These elasticities are estimated either by country-by-country regressions (6) with fitted β_i's represented by circles, or by panel regression (7) with fitted values represented by the solid line (red in online version).

Our strategy is to exploit this heterogeneity to more precisely estimate how economies respond to fluctuation in oil prices. Indeed, figure 6 illustrates why a country's oil dependence may go a long way in shaping the link between changes in oil prices and changes in consumption expenditures.

First, we run country-specific ordinary least squares regressions of consumption ($c_{i,t}$) on its lag, lagged real oil prices (oil_{t-1}), and world GDP (Y_t):

$$c_{i,t} = \rho_i c_{i,t-1} + \beta_i (1 - \rho_i) oil_{t-1} + \eta_i Y_t + u_{i,t}, \tag{6}$$

where the country-specific coefficient β_i can be interpreted as the long-run elasticity of consumption to changes in oil prices.[6] If all countries were to respond similarly to changes in oil prices, one should not find a relationship between elasticities β_i's and countries' oil dependence. Yet, as shown in figure 6, the elasticity of consumption to oil prices (β_i's in circles) is negatively correlated with oil dependency across countries.

Then, we propose a parsimonious approach to capture the heterogeneous responses of consumption to oil prices seen above. We estimate a panel regression in which the long-run elasticity of consumption to changes in oil prices, $\tilde{\beta}_{i,t}$, depends on the country's oil dependency, $d_{i,t}$, after a flexible transformation $g\left(d_{i,t}\right)$:

$$c_{i,t} = \rho c_{i,t-1} + \tilde{\beta}_{i,t}\left(1-\rho\right) oil_{t-1} + \eta Y_t + u_{i,t}, \tag{7}$$

$$\tilde{\beta}_{i,t} = b_0 + b_1 g(d_{i,t}; b_2, b_3). \tag{8}$$

Specifically, we assume that function $g\left(d_{i,t}\right)$ follows a strictly increasing Gompertz transformation:

$$g(d_{i,t}; b_2, b_3) = \exp\left\{-\exp\left[-b_2\left(d_{i,t}-b_3\right)\right]\right\}, \qquad b_2 > 0, \qquad b_3 \in \mathbb{R}. \tag{9}$$

By using this transformation, we can capture a wide range of relationships between an economy's oil dependency, $d_{i,t}$, and its long-run elasticity of consumption to oil prices, $\tilde{\beta}_{i,t}$: from a linear one to a nonlinear one in which lower oil dependency yields less than proportional increases in the elasticity to oil prices.[7] Indeed, the solid line (red in online version) of figure 6 shows that as countries become less oil dependent, the elasticity of consumption to oil prices predicted by regression (7) increases by a magnitude less than proportional to what would be predicted by a simple linear relationship.

[6]We take logs of all variables in this regression and detrend them with a cubic trend, with the exact detrending mattering little for the results. Also, the actual regression has two lags on consumption.

[7]We estimate the model (7)–(9) by nonlinear least squares. Consistent with figure 6, we also impose that $b_0 > 0$ and $b_1 < 0$.

3.2 Cross-Country Panel VAR

Motivated by the cross-country heterogeneity documented in the previous section, we estimate a hybrid panel VAR that enriches the benchmark VAR of section 2 by controlling for how the degree of oil dependency affects the macroeconomic response to an oil price shock. The data used are the same as in section 2: quarterly data for the 1975–2018 period, covering 55 countries, and totaling more than 6,700 observations. We specify the hybrid panel VAR as follows:

$$\begin{aligned} y_{i,t} &= \gamma_{yy} y_{i,t-1} + \gamma_{yY} Y_{t-1} + \gamma_{yc} c_{i,t-1} + \gamma_{yC} C_{t-1} + \gamma_{yz} z_{t-1} \\ &\quad + \left(\gamma_{yo} + \gamma_{yd} g\left(d_{i,t-1}\right)\right) oil_{t-1} \\ &\quad + \left(\delta_{yo} + \delta_{yd} g\left(d_{i,t-1}\right)\right) noil_{t-1} + \varepsilon^{y}_{i,t}, \qquad (10) \\ c_{i,t} &= \alpha_{cy} y_{i,t} + \alpha_{cY} Y_t + \gamma_{cy} y_{i,t-1} + \gamma_{cY} Y_{t-1} + \gamma_{cc} c_{i,t-1} + \gamma_{cC} C_{t-1} \\ &\quad + \gamma_{cz} z_{t-1} + \left(\gamma_{co} + \gamma_{cd} g\left(d_{i,t-1}\right)\right) oil_{t-1} \\ &\quad + \left(\delta_{co} + \delta_{cd} g\left(d_{i,t-1}\right)\right) noil_{t-1} + \varepsilon^{c}_{i,t}, \qquad (11) \\ z_t &= \alpha_{zY} Y_t + \alpha_{zC} C_t + \gamma_{zY} Y_{t-1} + \gamma_{zC} C_{t-1} + \gamma_{zz} z_{t-1} \\ &\quad + \gamma_{zo} oil_{t-1} + \varepsilon^{z}_{t}, \qquad (12) \\ oil_t &= \alpha_{oY} Y_t + \alpha_{oC} C_t + \alpha_{oz} z_t + \gamma_{oY} Y_{t-1} + \gamma_{oC} C_{t-1} + \gamma_{oz} z_{t-1} \\ &\quad + \gamma_{oo} oil_{t-1} + \varepsilon^{oil}_{t}, \qquad (13) \end{aligned}$$

$$noil_t \equiv \max\left(0, oil_t - \max\left(oil_{t-1}, oil_{t-2}, oil_{t-3}, oil_{t-4}\right)\right), \qquad (14)$$

where z_t is the Conference Board's composite index of leading indicators (CLI) of business cycles conditions,[8] oil_t is the real price of oil, $c_{i,t}$ is country-specific consumption, $y_{i,t}$ is country-specific GDP, and $C_t \equiv \sum_{i=1}^{N} \omega_{i,t} \cdot c_{i,t}$ and $Y_t \equiv \sum_{i=1}^{N} \omega_{i,t} \cdot y_{i,t}$ are the corresponding world aggregates, with $\omega_{i,t}$ denoting country weights.[9] All variables are measured in percent deviation from a quadratic trend.

[8] See Camacho and Perez-Quiros (2002) for the CLI's performance in anticipating business cycles conditions. The VAR results are similar if we use instead the Conference Board's composite index of coincident indicators or global oil production.

[9] The estimated VAR has an additional lag for every lagged variable and intercepts. We omit these variables in equations (10)–(13) to save on notation.

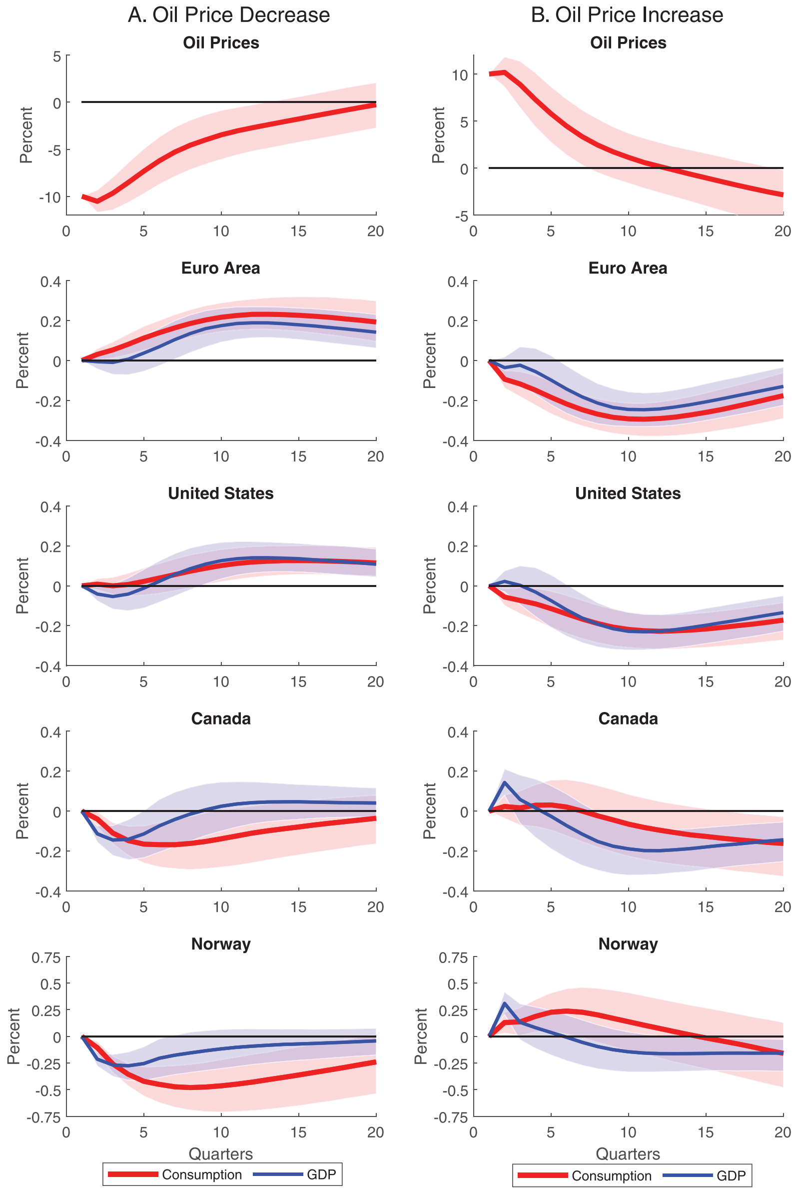
Note that this specification assumes that shocks to oil prices affect consumption, GDP, and commodity prices with a one-period delay.[10]

The specification above departs from a standard panel VAR (Holtz-Eakin, Newey, and Rosen 1988) in three ways. First, we allow countries' responses of consumption and GDP to oil price shocks to be a function of their time-varying oil dependence $d_{i,t}$. The specific functional form is the transformation $g(d_{i,t}; b_2, b_3)$ estimated in the panel model (7)–(9). Second, we allow individual countries' consumption and GDP to depend not only on their own lags but also on lags of the world aggregates, thus modeling dynamic interdependencies across countries. The world aggregates are averages of the country-specific variables, as done for instance in the global VAR approach proposed by Pesaran (2006) and discussed in Canova and Ciccarelli (2013) and Chudik and Pesaran (2016). Third, we augment the panel VAR with the net real oil price increase variable ($noil_t$), following Kilian and Vigfusson (2011) and similar to Hamilton (1996), thus allowing for increases and decreases in oil prices to yield asymmetric effects on economic activity, following the lead of Davis and Haltiwanger (2001) and Hamilton (2003).

Figure 7 plots the IRFs of consumption and GDP to oil price shocks using the 2018 values of countries' oil dependencies. The choice of countries showcases the heterogeneous responses arising from these differing oil dependencies. When oil prices decline (figure 7A), consumption and GDP of oil-importer countries, such as those in the euro area, increase. However, these increases take time to materialize, reaching peaks after roughly 12 quarters. As in section 2, consumption responds before GDP, suggesting that wealth effects are larger than supply-side effects. Conversely, consumption and GDP of oil-exporter countries, such as Canada and Norway, quickly decrease, bottoming out six and two quarters after the shock, respectively. These results are consistent with a rapid tumble in investment, presumably in the energy sector, which then compounds with a protracted fall in consumption, possibly due to

[10] To see this, notice that only the lags of oil prices (oil_{t-1}) and net oil price increases ($noil_{t-1}$) are present in equations (10) and (11). However, differently from a standard VAR, the cross-equations restrictions embedded in the panel VAR ($C_t \equiv \sum_{i=1}^{N} \omega_{i,t} \cdot c_{i,t}$ and $Y_t \equiv \sum_{i=1}^{N} \omega_{i,t} \cdot y_{i,t}$) produce errors that are generally correlated across equations. See Canova and Ciccarelli (2013) for discussion.

Figure 7. Impulse Response Functions to Oil Price Shocks to Cross-Country Panel VAR



Notes: The figure shows the IRFs to oil price shocks in the cross-country panel VAR (section 3.2). Solid lines represent median IRFs. Shaded areas represent 16th and 84th percentiles constructed with 5,000 bootstrap replications. Variables are expressed as percent deviation from quadratic trend.

Table 1. Responses of Consumption and GDP in Cross-Country Panel VAR

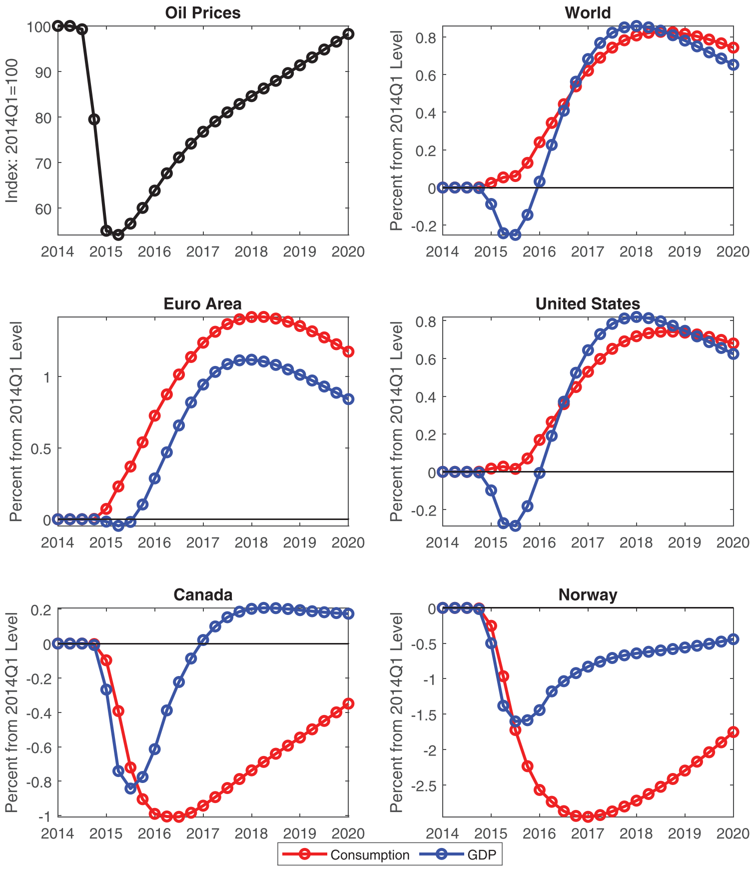
	Euro Area	United States	Canada	Norway	World
10% Oil Price Decline					
Consumption	0.16	0.06	−0.13	−0.40	0.07
GDP	0.10	0.06	−0.04	−0.17	0.06
10% Oil Price Increase					
Consumption	−0.23	−0.16	−0.02	0.16	−0.17
GDP	−0.16	−0.14	−0.09	−0.03	−0.14

Notes: The consumption and GDP responses are measured as the average of their impulse response functions during the first three years (from quarter 1 through quarter 12) after a shock moving oil prices by 10 percent (in quarter 0). These IRFs are calculated using the cross-country panel VAR from section 3.2.

negative wealth effects. The United States, neither a large net oil importer nor a large net oil exporter in 2018, experiences more mixed effects: its GDP leans towards a contraction in the periods immediately after the shock, and only rises above trend many quarters later, pulled by a modest rise in consumption.

Figure 7 also shows that oil price increases (figure 7B) do more harm than the good afforded by oil price decreases (figure 7A). In oil-importing economies, such as those in the euro area, the negative responses of GDP and consumption to oil price increases are larger in magnitude than the positive responses to oil price decreases, as in Hamilton (2003). In large oil-exporting economies, such as Norway, GDP and consumption experience short-lived expansions after an oil price increase, with these variables returning to their original level after six quarters. This result contrasts with the protracted drag to economic activity in the aftermath of oil price declines. One possible reason for such a small boost to oil-exporter economies from oil price increases is that these economies could suffer large drags from the downturn experienced by their oil-importer trading partners (which account for a much larger share of the world GDP). Table 1 summarizes these results, showing that the world GDP experiences only small boosts after oil price declines and much larger drags after oil price increases.

Figure 8. Contribution of the 2014–15 Oil Price Slump to GDP and Consumption



Notes: The figure shows the historical contribution of oil price shocks to consumption and GDP of the world aggregate and selected countries. Oil price shocks are those estimated for the periods 2014:Q3 through 2015:Q1 using the cross-country panel VAR (section 3.2). Variables are expressed as percent deviation from the 2014:Q1 level, with the exception of oil prices, which is measured as an index with level 100 at 2014:Q1.

In figure 8, we use the estimates of the cross-country panel VAR to quantify how the "exogenous" portion of the 2014–15 oil price slump affected consumption and GDP across countries. First and foremost, the VAR attributes a large chunk of the oil price decline throughout that period to surprise innovations to the oil price equation (ε_t^{oil} in equation (13)) in periods 2014:Q3–2015:Q1. Other shocks, as well as delayed effects from other shocks taking place before 2014:Q3, play a much more limited role in driving oil prices in 2014–15. Accordingly, the response of oil prices—shown in the top-left panel in terms of the 2014:Q1 level—closely mirrors the actual path of oil prices.

The remaining panels of figure 8 show the response of consumption and GDP across countries. The top-right panel shows the response for the world. Consumption rises gradually, peaking at about 0.8 percent above baseline in mid-2018. The response of GDP is initially slightly negative, likely reflecting some drag from declines in investment in the oil sector, but also builds to about 0.8 percent by the end of 2017. The middle panel compares the euro area and the United States. While the U.S. response mirrors that of the world as a whole, the response of the euro area to the oil price decline is larger than that of the United States, reflecting the euro zone's larger oil dependence. In the bottom panel, GDP and consumption fall markedly in Canada and Norway, consistent with the fact that both economies are large oil exporters.

In appendix B, we conduct two robustness exercises. First, we compare the oil price shocks ε_t^{oil} estimated by our cross-country panel VAR (equations (10)–(14)) with the oil-specific demand shocks identified by Kilian (2009). Figure B.1 shows that these shocks are similar, exhibiting a correlation of 0.49, which is higher than the one found by the analogous comparison done in section 2.2. Second, we estimate our panel VAR starting in 1986, given a possible structural break in oil dynamics around this period (e.g., Baumeister and Peersman 2013). The impulse responses for the restricted sample, shown in figure B.2, are similar to those from our baseline model.

4. Oil Dependency across U.S. States: State-Level Panel VAR

In this section, we analyze how economic activity reacts to oil price shocks across U.S. states by exploiting the cross-section and time-series variation in states' oil dependence. We show that consumption (proxied by car registrations) of states with higher oil dependency increases more after exogenous oil price declines. We also show that changes in oil prices generate asymmetric effects: oil price decreases generate smaller boosts to aggregate U.S. consumption than the drag from oil price increases.

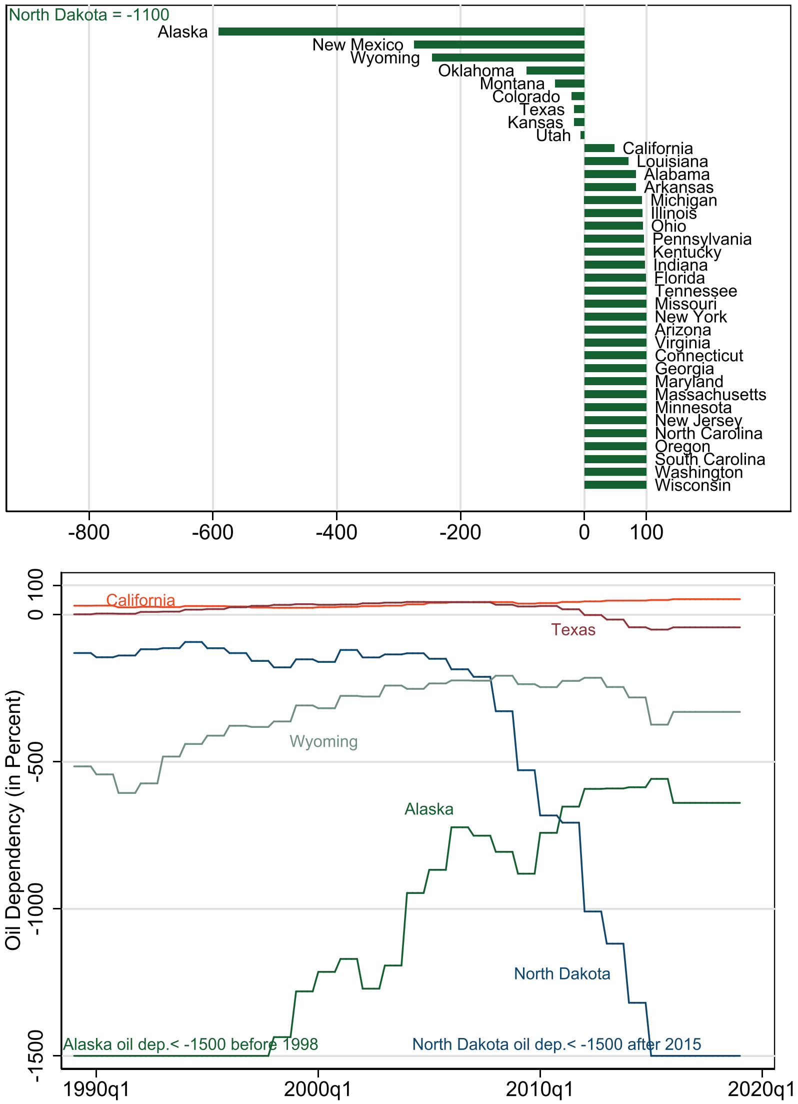
4.1 Oil Dependency across the United States

Just like countries in the world, U.S. states are very heterogeneous in their oil dependence.[11] Figure 9 provides information about cross-sectional and time-series variation in U.S. states' oil dependency, using the measure defined in equation (5). The top panel of figure 9 shows that in 2013 states in the Northeast had an oil dependence of nearly 100 percent, while Texas was roughly oil independent. By contrast, states such as Alaska and North Dakota produced substantially more oil than consumed, with very large negative oil dependencies. The bottom panel of figure 9 shows that U.S. states also exhibit considerable time-series variation in oil dependence, with North Dakota exporting increasing amounts of oil over time and Alaska exporting decreasing amounts. Table A.2 in appendix A provides more details on oil dependency across U.S. states.

In order to estimate the response of consumption to oil price shocks at the state level, we use data on retail new car registrations as a proxy for consumption. Available at a quarterly frequency from 1989:Q1 to 2018:Q4, this data set constructed by Polk's National Vehicle Population Profile has important advantages over state-level data on personal consumption expenditures, as the latter are still experimental, are available only annually, and cover a shorter time frame. Appendix A describes the state-level car registrations data and shows that retail registrations at the national level are highly correlated with NIPA (national income and product accounts) expenditures on motor vehicles and parts.

[11] In fact, using the cross-sectional standard deviation of oil dependency as a measure of heterogeneity, the U.S. states are actually more heterogeneous than the world's countries in terms of oil dependency. In 2013, the cross-sectional standard deviation of oil dependency across U.S. states was 200. The corresponding number for the world's countries was 130.

Figure 9. Oil Dependency across U.S. States and over Time



Notes: The top panel shows oil dependency in 2013 for a subset of the largest states in the United States (oil producers and states with share of car purchases in 2013 larger than 1 percent), where oil dependency is measured by oil imports as a percentage of oil consumption (equation (5)). The bottom panel plots oil dependency over time for a select group of U.S. states.

4.2 State-Level Panel VAR

We set up a panel VAR that mirrors the cross-country analysis discussed in the previous section. To do so, we use unemployment rates to measure overall economic activity at the state level. Thus, for each state i, let $u_{i,t}$ denote the unemployment rate in quarter t and $car_{i,t}$ the car registrations, with $U_t \equiv \sum_{i=1}^{N} \omega_{i,t} u_{i,t}$ referring to the national unemployment, $Car_t \equiv \sum_{i=1}^{N} \omega_{i,t} car_{i,t}$ to the national car registrations, and $\omega_{i,t}$ to state-specific weights. The rest of the notation is the same as in section 3: oil_t is the real price of oil, $noil_t$ is the real net oil price increase (equation (14)), and z_t is the CLI. All variables are measured in percent deviation from a quadratic trend, with the exception of unemployment rates, which are measured as percentage-point deviation from a quadratic trend. With a specification analogous to the one of section 3.2, the state-level panel VAR allows for heterogeneous regional responses to an oil shock as a function of oil dependence:

$$\begin{aligned} u_{i,t} = {} & \gamma_{uu} u_{i,t-1} + \gamma_{uU} U_{t-1} + \gamma_{uc} car_{i,t-1} + \gamma_{uC} Car_{t-1} \\ & + \gamma_{uz} z_{t-1} + \left(\gamma_{uo} + \gamma_{ud} g\left(d_{i,t-1}\right)\right) oil_{t-1} \\ & + \left(\delta_{uo} + \delta_{ud} g\left(d_{i,t-1}\right)\right) noil_{t-1} + \varepsilon_{i,t}^{u}, \end{aligned} \tag{15}$$

$$\begin{aligned} car_{i,t} = {} & \alpha_{cu} u_{i,t} + \alpha_{cU} U_t + \gamma_{cu} u_{i,t-1} + \gamma_{cU} U_{t-1} + \gamma_{cc} car_{i,t-1} \\ & + \gamma_{cC} Car_{t-1} + \gamma_{cz} z_{t-1} + \left(\gamma_{co} + \gamma_{cd} g\left(d_{i,t-1}\right)\right) oil_{t-1} \\ & + \left(\delta_{co} + \delta_{cd} g\left(d_{i,t-1}\right)\right) noil_{t-1} + \varepsilon_{i,t}^{c}, \end{aligned} \tag{16}$$

$$\begin{aligned} z_t = {} & \alpha_{zU} U_t + \alpha_{zC} Car_t + \gamma_{zU} U_{t-1} + \gamma_{zC} Car_{t-1} + \gamma_{zz} z_{t-1} \\ & + \gamma_{zo} oil_{t-1} + \varepsilon_{t}^{z}, \end{aligned} \tag{17}$$

$$\begin{aligned} oil_t = {} & \alpha_{oU} U_t + \alpha_{oC} Car_t + \alpha_{oz} z_t + \gamma_{oU} U_{t-1} + \gamma_{oC} Car_{t-1} \\ & + \gamma_{oz} z_{t-1} + \gamma_{oo} oil_{t-1} + \varepsilon_{t}^{oil}, \end{aligned} \tag{18}$$

where the functional form of $g\left(d_{i,t}\right)$ is described in equation (9) with coefficients estimated by a regression analogous to (7).[12]

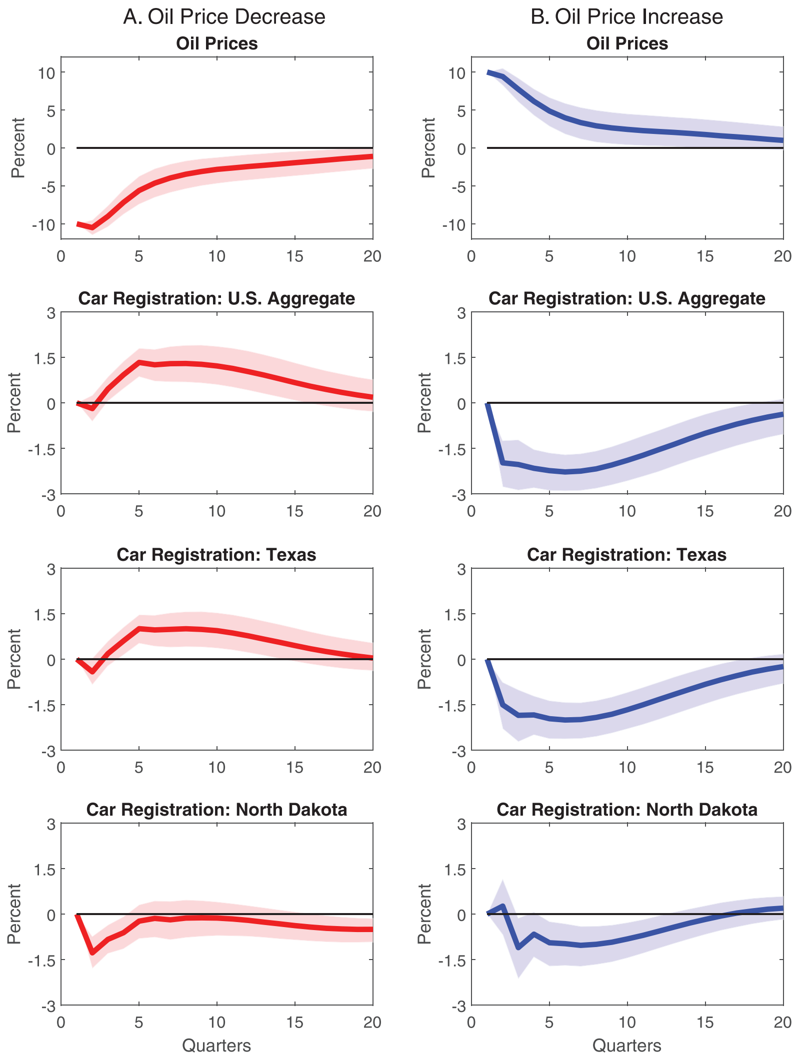
[12]The estimated VAR has an additional lag for every lagged variable and intercepts. We omit these variables in equations (15)–(18) to save on notation.

Figure 10A shows the predicted effect on car registrations of a 10 percent oil price decline triggered by an ε_t^{oil} shock. For the United States in aggregate, such shock is associated with a substantial increase in car registrations, peaking at 1.3 percent a year after the shock, sustaining this increase for another year, and then experiencing fading effects. The large response of car registrations is consistent with the channels that emphasize the large oil price elasticity of demand for goods that are complementary with oil, and lines up well with the findings in the literature. For example, Edelstein and Kilian (2009) find that vehicle expenditures are highly sensitive to movements in the price of gasoline. Coglianese et al. (2017) find an estimate of −0.37 for the price elasticity of gasoline demand, although their numbers do not appear directly comparable to ours, as they look at the cumulative effects of gasoline price changes.

We then turn to the state-specific responses of car registrations to oil price drops (figure 10A). Oil-importing states face effects similar to those seen for the United States in the aggregate, and thus are omitted from the figure. Oil-independent Texas has a response similar to the U.S. aggregate, but with a magnitude slightly smaller. By contrast, major oil-exporting states such as North Dakota suffer a substantial decline in car registrations, with a trough of −1.3 percent in the quarter right after the shock. North Dakota's car registrations soon rise, but only to a level not statistically different from zero. This rise may reflect the fact that, because cars run on gas, lower oil prices make car-buying more attractive, even in oil-intensive states. Moreover, the subsequent return of car registrations to North Dakota's original level may also reflect the increase in overall U.S. aggregate demand, with the close trade and financial linkages within the U.S. states dampening the effects of heterogeneity in oil production.

Figure 10B then shows that the benefits from oil price decreases are smaller than the losses from oil price increases also across U.S. states. We see that shocks pushing up oil prices are followed by a national decrease in car registrations. This decrease is larger than the increase in registrations following a comparable oil price drop (figure 10A). The results for Texas are qualitatively similar to those for the national aggregate, although with smaller magnitudes. In the case of oil-rich North Dakota, car registrations also fall after an oil price increase, possibly due to the U.S. aggregate demand effect. Figure B.4 in the appendix with supplementary results (appendix B)

Figure 10. Impulse Response Functions to Oil Price Shocks in State-Level Panel VAR



Notes: The figure shows the IRFs to oil price shocks in the U.S. state-level panel VAR (section 4). Solid lines represent median IRFs. Shaded areas represent 16th and 84th percentiles constructed using 5,000 bootstrap replications. Variables are expressed as percent deviation from quadratic trend.

shows that results for unemployment rates are similar to those for car registrations, with the main difference being that unemployment rates react a bit more slowly to oil price shocks than car registrations.

All told, the state-level evidence supports the findings of the country-level VAR. It highlights once again how varying oil dependencies may imply differences in consumption responses across regions, even if these regions face common monetary and fiscal policies. In fact, as is well known in the literature on fiscal and monetary unions, a common policy response may either amplify or dampen the heterogeneity of regional responses, especially when shocks present authorities with stabilization tradeoffs. Consider for instance an oil shock that lowers inflation and activity in North Dakota but lowers inflation and boosts activity in Connecticut. If the policy goal is to stabilize activity, such shock calls for expansionary policies in North Dakota and for contractionary policies in Connecticut. Thus, heterogeneous policy responses would imply responses of activity in North Dakota and Connecticut being more similar than under the case of a common policy response.

5. Literature Review

There is a vast literature that analyzes how fluctuations in oil prices affect macroeconomic outcomes. In particular, our paper belongs to a growing set of studies that investigate the responses to oil price shocks across different economies. These studies have explored variation in geographical location, institutional development, and net-exporting position of energy across different countries and states of the Unites States (e.g., Aastveit, Bjørnland, and Thorsrud 2015; Bjørnland and Zhulanova 2018; Guerrero-Escobar, Hernandez-del-Valle, and Hernandez-Vega 2019; and Peersman and Van Robays 2012). In addition, Baumeister, Peersman, and Van Robays (2010) (henceforth abbreviated as BPVR) is particularly related to our work because of its cross-country perspective. BPVR estimate country-specific VARs for eight different countries and document that net-energy importing economies suffer declines in economic activity following an oil supply shock, with these effects being not significant or positive for net-energy exporters. BPVR also show qualitative evidence that improvements in the net-energy position of countries are associated with more muted responses to oil supply shocks.

Our paper makes four contributions. First, we explore a much larger set of economies than previous papers by using two panel data sets: one with 55 countries and another with all U.S. states. Second, we exploit the variation in oil dependency both across countries and over time—beyond the status of oil importer/exporter—thus better quantifying how changes in oil dependency affect an economy's response to oil price shocks. Third, we measure the effect of oil shocks on aggregate consumption, a variable often overlooked in the literature. We show that oil price declines generate large and positive effects on consumption and economic activity (GDP and unemployment rate) in oil-importing economies, while depressing consumption and activity in oil-exporting economies. Fourth, we allow asymmetric responses to oil price shocks, with macroeconomic effects from oil price increases being potentially different from oil price decreases. We then show that oil price increases do more harm than the good afforded by oil price decreases across both the world and the United States.

6. Conclusions

In this paper, we provide empirical evidence that lower oil prices boost consumption in oil-importing countries and depress it in oil-exporting ones. This effect varies in proportion to the degree of oil dependency of an economy. For instance, a 10 percent decline in oil prices boosts euro-area consumption and GDP by about 0.2 percent, while leading to declines of similar magnitudes in consumption and GDP in oil-exporting countries such as Canada. In between, there are countries like the United States where the effects are more mixed. While in the short run U.S. GDP might temporarily decline, presumably reflecting weaker investment in the oil industry, a gradual rise in U.S. consumption pushes its GDP toward a modest expansion. For the world aggregate, oil price declines boost economic activity, although by smaller magnitudes than similar increases in oil prices. We complement our cross-country results by showing analogous findings for the U.S. states and the U.S. aggregate.

Our results have important implications for the design of monetary policy in the presence of oil price fluctuations. A large and growing literature documents how optimal monetary policy may entail different responses to oil price increases depending on the source of the shock, say an increase in foreign demand or a glut in foreign supply. For instance, Bodenstein, Guerrieri, and Kilian (2012) develop and estimate a dynamic stochastic general equilibrium model with oil in which multiple shocks drive oil price fluctuations. In their model, the peak response of output and consumption to oil shocks happens in the same period of the shock. They find that no two shocks induce the same policy response, even controlling for the same path of oil prices implied by the shocks. They further show that, in the wake of fluctuations in oil prices, a monetary policy rule that places a higher weight on stabilizing wage inflation fosters the stabilization of core inflation. Our evidence reinforces their argument, by showing that the same shock may imply different macroeconomic responses depending on the oil dependency of a particular country. It also illustrates an important challenge for extant models of the oil market and the macroeconomy, which may want to incorporate features such as habits and adjustment costs in order to capture the delays in which oil shocks affect economic activity.

Our evidence shows that changes in oil dependence may imply nontrivial effects of changes in oil prices, that the effects of oil shocks on consumption may take time to materialize, and that the response of GDP to an oil shock is not equal to the mere contribution of consumption, presumably because investment in the energy sector may be sensitive to changes in oil prices too. We view such evidence, gathered from a large set of countries over a long time period, as providing a useful empirical benchmark to gauge the plausibility of the policy recommendations of optimizing monetary models of the business cycle that incorporate an oil sector.

Appendix A. The Data

A.1 Country-Level Data

The sample includes 55 countries for which it was possible to assemble sufficiently long and reliable data on quarterly real consumption and quarterly real GDP from the country's official statistics. Of the richest countries in the world (measured by total GDP in U.S. dollars), we do not include China because of the lack of consumption data at a quarterly frequency. By the same reason, we exclude some of the largest oil producers, such as United Arab Emirates, Iraq, Kuwait, Nigeria, Qatar, Angola, Algeria, and Kazakhstan.

In the aggregate VAR analysis, we construct GDP and consumption using data for real consumption and real GDP using weights based on each country's GDP in constant U.S. dollars from the World Bank World Development Indicators (if the data are missing, the weight is set at zero, and the weights of other countries are changed accordingly). We list in table A.1 all countries in our sample, their oil dependency, their sample weights for the year 2013, and their data coverage.

Oil dependency is constructed using annual data on oil production and consumption from the BP Statistical Review of World Energy (with data dating back to 1965). An advantage of these data is that, barring small changes in year-to-year inventories, they satisfy the identity that oil production must be equal oil consumption over very long horizons. For countries that are small oil producers, the BP Statistical Review does not report data (shown as missing in table A.1). In these cases, we search for oil production and consumption in the CIA World Factbook and the International Energy Agency (IEA). If we find these oil statistics, we set oil dependency to a constant implied by the data for all years. If we are not able to find the data in either the BP Statistical Review, the CIA World Factbook, or the IEA, we set oil dependency to 1 in all years. In all these cases of small oil producers, the implied oil dependency is above 0.65.

A.2 State-Level Data

The key variable included in our state-level panel vector autoregression is "new car registrations, retail" broken down by state of registration. This measure is constructed using data from Polk's National Vehicle Population Profile, which is a census of all currently registered passenger cars and light-duty trucks in the United States. Polk's New Registration Data provides detailed indicators for new vehicle registrations (including where the vehicle was registered). The "retail" file represents new vehicles purchased by individuals,

Table A.1. Cross-Country Oil Dependency in 2013 and Data Availability

Country	GDP Share	Oil Dependency	Oil Consumption	Oil Production	Exporter	Start Date	End Date
United States	27.03	47	18,961	10,071	0	1975:Q1	2018:Q4
Japan	10.05	97	4,516		0	1975:Q1	2018:Q4
Germany	6.10	93	2,408		0	1975:Q1	2018:Q4
France	4.65	97	1.664		0	1975:Q1	2018:Q4
United Kingdom	4.40	43	1,518	864	0	1975:Q1	2018:Q4
Brazil	4.11	32	3,124	2,110	0	1990:Q1	2018:Q4
Italy	3.48	91	1,260	114	0	1975:Q1	2018:Q4
India	3.37	76	3,727	906	0	1996:Q2	2018:Q4
Canada	2.96	−68	2,383	4,000	1	1975:Q1	2018:Q4
Russia	2.89	−245	3,135	10,809	1	1995:Q1	2018:Q4
Spain	2.31	97	1,195		0	1975:Q1	2018:Q4
Australia	2.13	61	1,034	407	0	1975:Q1	2018:Q4
Korea	2.04	96	2,455		0	1975:Q1	2018:Q4
Mexico	1.96	−41	2,034	2,875	1	1980:Q1	2018:Q4
Turkey	1.66	100	757		0	1987:Q1	2018:Q4
Indonesia	1.53	47	1,663	882	0	1983:Q1	2018:Q4
Netherlands	1.45	94	898		0	1975:Q1	2018:Q4
Saudi Arabia	1.07	−230	3,451	11,393	1	2005:Q1	2018:Q4
Switzerland	1.04	100	249		0	1975:Q1	2018:Q4
Poland	0.88	93	538		0	1995:Q1	2018:Q4
Sweden	0.86	96	309		0	1975:Q1	2018:Q4
Belgium	0.84	98	636		0	1975:Q1	2018:Q4

(continued)

Table A.1. (Continued)

Country	GDP Share	Oil Dependency	Oil Consumption	Oil Production	Exporter	Start Date	End Date
Taiwan	0.83	100	1,010		0	1975:Q1	2018:Q4
Iran	0.79	−80	2,011	3,617	1	2004:Q2	2018:Q2
Argentina	0.78	5	683	647	0	1993:Q1	2018:Q4
Norway	0.77	−656	243	1,838	1	1975:Q1	2018:Q4
Venezuela	0.75	−243	782	2,680	1	1997:Q1	2015:Q4
Austria	0.69	87	264		0	1975:Q1	2018:Q4
South Africa	0.69	73	572		0	1975:Q1	2018:Q4
Thailand	0.65	65	1,299	452	0	1993:Q1	2018:Q4
Colombia	0.57	−237	298	1,004	1	2000:Q1	2018:Q4
Denmark	0.56	−13	158	178	1	1975:Q1	2018:Q4
Malaysia	0.51	22	803	626	0	1991:Q1	2018:Q4
Singapore	0.47	98	1,225		0	1975:Q1	2018:Q4
Israel	0.44	98	247		0	1995:Q1	2018:Q4
Chile	0.43	95	362		0	1996:Q1	2018:Q4
Hong Kong	0.43	100	311		0	1990:Q1	2018:Q4
Finland	0.42	95	191		0	1975:Q1	2018:Q4
Greece	0.42	98	295		0	1975:Q1	2018:Q4
Philippines	0.40	94	326		0	1981:Q1	2018:Q4
Ireland	0.40	99	138		0	1975:Q1	2018:Q4
Portugal	0.38	97	239		0	1975:Q1	2018:Q4
Czech Republic	0.36	94	184		0	1996:Q1	2018:Q4
Peru	0.30	26	227	167	0	1980:Q1	2018:Q4
New Zealand	0.27	66	151		0	1975:Q1	2018:Q4
Hungary	0.23	82	129		0	1995:Q1	2018:Q4

(*continued*)

Table A.1. (Continued)

Country	GDP Share	Oil Dependency	Oil Consumption	Oil Production	Exporter	Start Date	End Date
Slovakia	0.16	89	75		0	1995:Q1	2018:Q4
Ecuador	0.14	−113	247	527	1	1990:Q1	2018:Q4
Luxembourg	0.10	100	57		0	1975:Q1	2018:Q4
Slovenia	0.08	99	50		0	1995:Q1	2018:Q4
Latvia	0.05	97	33		0	1995:Q1	2018:Q4
Estonia	0.04	65	31		0	1995:Q1	2018:Q4
El Salvador	0.03	100	25		0	1990:Q1	2018:Q4
Botswana	0.03	100	17		0	1994:Q1	2018:Q4
Iceland	0.03	100	15		0	1997:Q1	2018:Q4

Notes: Oil dependency is calculated as oil consumed minus oil produced as a percentage of oil consumed (source: BP Statistical Review of World Energy, downloaded in June 2019). Oil production is production of crude oil, tight oil, oil sands, and natural gas liquids (NGLs) (thousand barrels daily). Oil consumption is inland demand plus international aviation and marine bunkers and refinery fuel and loss (thousand barrels daily). In the aggregate VAR, the "exporters" are countries with negative oil dependency in 2013.

Table A.2. U.S. State-Level Data in 2013

State	Car Share	Oil Dependency	Oil Consumption	Oil Production
California	11.45	48	1,713	891
Texas	8.55	−16	3,564	4,153
Florida	7.14	99	824	10
New York	5.77	100	658	2
Pennsylvania	4.45	96	618	24
Illinois	4.26	93	627	43
Michigan	4.21	93	449	32
Ohio	4.17	94	589	36
New Jersey	3.55	100	516	0
Georgia	2.94	100	471	0
North Carolina	2.86	100	437	0
Virginia	2.58	100	425	0
Massachusetts	2.25	100	298	0
Maryland	2.17	100	253	0
Arizona	1.88	100	267	0
Indiana	1.86	97	398	11
Missouri	1.86	100	325	1
Tennessee	1.84	100	352	1
Wisconsin	1.81	100	279	0
Washington	1.71	100	369	0
Louisiana	1.62	71	1,102	322
Colorado	1.48	−21	246	297
Alabama	1.48	83	268	47
Minnesota	1.47	100	318	0
South Carolina	1.36	100	262	0
Connecticut	1.27	100	167	0
Oklahoma	1.14	−93	267	516
Kentucky	1.10	97	303	10
Oregon	1.09	100	172	0
Arkansas	0.94	83	172	30
Iowa	0.92	100	241	0
Kansas	0.79	−16	180	210
Mississippi	0.79	50	219	109
Nevada	0.77	99	119	1
West Virginia	0.65	69	106	32
Utah	0.63	−6	148	157
New Hampshire	0.60	100	80	0
New Mexico	0.60	−275	123	460
Nebraska	0.55	90	123	13
Maine	0.44	100	99	0
Hawaii	0.40	100	116	0
Delaware	0.36	100	51	0

(*continued*)

Table A.2. (Continued)

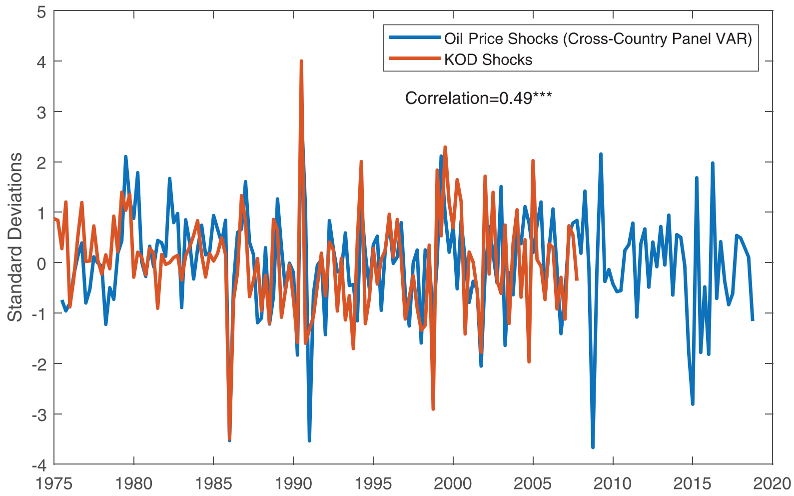
State	Car Share	Oil Dependency	Oil Consumption	Oil Production
Idaho	0.36	100	84	0
Rhode Island	0.33	100	44	0
Vermont	0.28	100	44	0
Montana	0.28	−47	89	131
South Dakota	0.24	86	61	8
North Dakota	0.22	−1,119	115	1,398
Alaska	0.20	−591	122	841
Wyoming	0.19	−246	82	284
District of Columbia	0.14	100	10	0

Notes: Oil dependency is calculated as oil consumed minus oil produced as a percentage of oil consumed (source: U.S. Energy Information Administration: State Profiles and Energy Estimates). Oil production is production of crude oil by state (*sumcrudestate*). Production of crude oil by state is scaled up by the ratio of total U.S. oil production from the BP Statistical Review (which includes tight oil, oil sands, and other NGLs and including offshore production) to *sumcrudestate* (which is about 1.5), so as to facilitate comparison with oil-dependency measures by country. Oil consumption includes all petroleum products consumed, and lines up with oil consumption from the BP Statistical Review. Production and consumption are in thousands of barrels daily.

thus excluding registrations of rental, fleet, government, and other commercial use vehicles. From 1989 to 2002, Polk's data is computed by marketing regions, with information also available for states. We then splice this data set with the one for 2002 onward, which is computed directly at the state level. As shown by figure B.3, retail registrations at the national level are very highly correlated with durable goods, namely motor vehicles and parts from the NIPA.

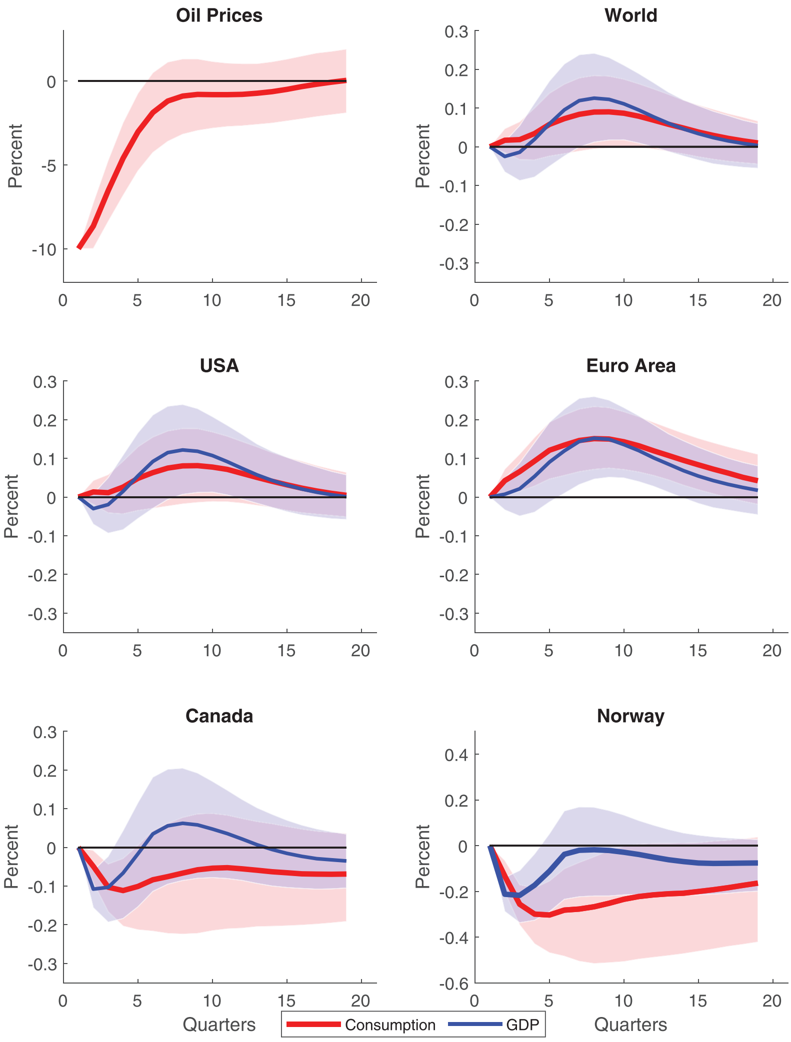
Appendix B. Supplementary Results

Figure B.1. Comparison of Oil Shocks Identified under Different Strategies



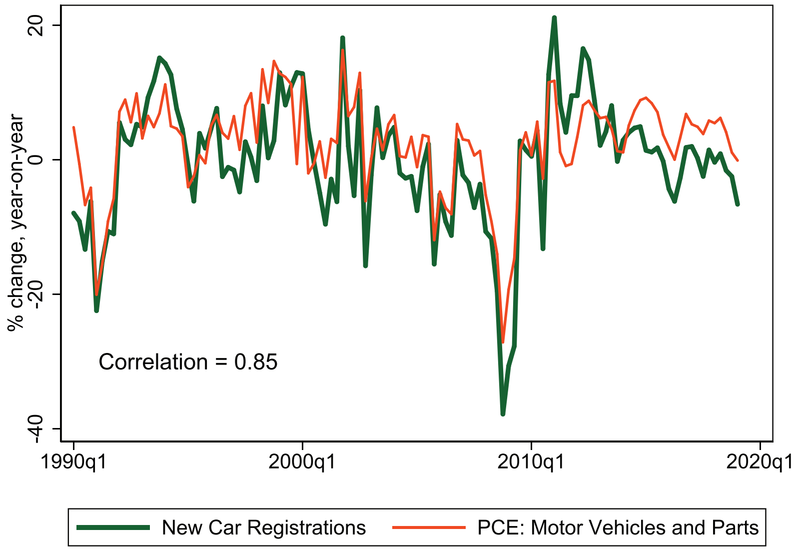
Notes: "Oil Price Shocks (Cross-Country Panel VAR)" represent the oil shocks identified by the panel VAR from section 3.2 of this paper. "KOD Shocks" represent the "oil-specific demand" shocks identified by Kilian (2009) averaged to the quarterly frequency. The correlation between the two series of shocks is statistically significant at the 99 percent confidence level (***).

Figure B.2. Impulse Response Functions to Oil Price Shocks in Cross-Country Panel VAR, Data Post-1986



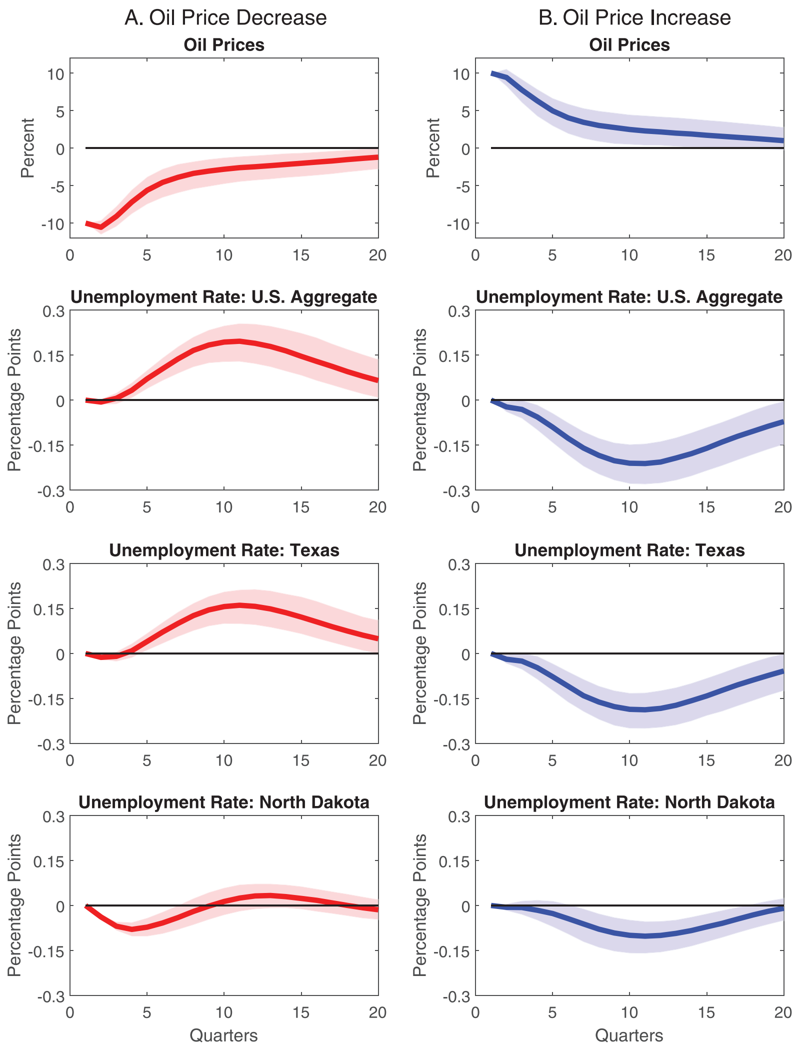
Notes: The figure shows the IRFs to oil price shocks in the cross-country panel VAR (section 3.2) using data from 1986 onward. Solid lines represent median IRFs. Shaded areas represent 16th and 84th percentiles constructed with 5,000 bootstrap replications. Variables are expressed as percent deviation from quadratic trend.

Figure B.3. New Cars' Registrations and Motor Vehicles Expenditures from NIPA



Note: This figure compares NIPA expenditures on motor vehicles (in real terms) with the aggregate data on new cars' registrations from Polk's National Vehicle Population Profile.

Figure B.4. Impulse Response Functions to Oil Price Shocks in State-Level Panel VAR



Notes: This figure shows the IRFs to oil price shocks in the U.S. state-level panel VAR (section 4). Solid lines represent median IRFs. Shaded areas represent 16th and 84th percentiles constructed using 5,000 bootstrap replications. Variables are expressed as percent deviation from quadratic trend.

References

Aastveit, K. A., H. C. Bjørnland, and L. A. Thorsrud. 2015. "What Drives Oil Prices? Emerging Versus Developed Economies." *Journal of Applied Econometrics* 30 (7): 1013–28.

Arezki, R., and O. Blanchard. 2014. "Seven Questions about the Recent Oil Price Slump." IMF Blog, December 22.

Baumeister, C., and G. Peersman. 2013. "The Role of Time-Varying Price Elasticities in Accounting for Volatility Changes in the Crude Oil Market." *Journal of Applied Econometrics* 28 (7): 1087–1109.

Baumeister, C., G. Peersman, and I. Van Robays. 2010. "The Economic Consequences of Oil Shocks: Differences across Countries and Time." In *Inflation in an Era of Relative Price Shock*, 91–128. Reserve Bank of Australia.

Bernanke, B. S. 2016. "The Relationship between Stocks and Oil Prices." Brookings Institution.

Bernanke, B. S., J. Boivin, and P. Eliasz. 2005. "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach." *Quarterly Journal of Economics* 120 (1): 387–422.

Bjørnland, H. C., and J. Zhulanova. 2018. "The Shale Oil Boom and the US Economy: Spillovers and Time-Varying Effects." Working Paper No. 8/2018, Centre for Applied Macro- and Petroleum Economics (CAMP), BI Norwegian Business School.

Bodenstein, M., C. J. Erceg, and L. Guerrieri. 2011. "Oil Shocks and External Adjustment." *Journal of International Economics* 83 (2): 168–84.

Bodenstein, M., L. Guerrieri, and L. Kilian. 2012. "Monetary Policy Responses to Oil Price Fluctuations." *IMF Economic Review* 60 (4): 470–504.

Caldara, D., M. Cavallo, and M. Iacoviello. 2019. "Oil Price Elasticities and Oil Price Fluctuations." *Journal of Monetary Economics* 103 (May): 1– 20.

Camacho, M., and G. Perez-Quiros. 2002. "This Is What the Leading Indicators Lead." *Journal of Applied Econometrics* 17 (1): 61–80.

Canova, F., and M. Ciccarelli. 2013. "Panel Vector Autoregressive Models: A Survey." ECB Working Paper No. 1507.

CBS/AP. 2016. "Buffett Says U.S. Economy Weaker Than He Expected but Growing." CBS MoneyWatch, February 29.

Chudik, A., and M. H. Pesaran. 2016. "Theory and Practice of GVAR Modelling." *Journal of Economic Surveys* 30 (1): 165–97.

Coglianese, J., L. W. Davis, L. Kilian, and J. H. Stock. 2017. "Anticipation, Tax Avoidance, and the Price Elasticity of Gasoline Demand." *Journal of Applied Econometrics* 32 (1): 1–15.

Davis, S. J., and J. Haltiwanger. 2001. "Sectoral Job Creation and Destruction Responses to Oil Price Changes." *Journal of Monetary Economics* 48 (3): 465–512.

Edelstein, P., and L. Kilian. 2009. "How Sensitive Are Consumer Expenditures to Retail Energy Prices?" *Journal of Monetary Economics* 56 (6): 766–79.

Guerrero-Escobar, S., G. Hernandez-del-Valle, and M. Hernandez-Vega. 2019. "Do Heterogeneous Countries Respond Differently to Oil Price Shocks?" *Journal of Commodity Markets* 16 (December).

Hamilton, J. D. 1996. "This Is What Happened to the Oil Price-Macroeconomy Relationship." *Journal of Monetary Economics* 38 (2): 215–20.

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

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

Holtz-Eakin, D., W. Newey, and H. S. Rosen. 1988. "Estimating Vector Autoregressions with Panel Data." *Econometrica* 56 (6): 1371–95.

Jiménez-Rodríguez, R., and M. Sánchez. 2005. "Oil Price Shocks and Real GDP Growth: Empirical Evidence for Some OECD Countries." *Applied Economics* 37 (2): 201–28.

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

Kilian, L., A. Rebucci, and N. Spatafora. 2009. "Oil Shocks and External Balances." *Journal of International Economics* 77 (2): 181–94.

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.

Peersman, G., and I. Van Robays. 2012. "Cross-country Differences in the Effects of Oil Shocks." *Energy Economics* 34 (5): 1532–47.

Pesaran, M. H. 2006. "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure." *Econometrica* 74 (4): 967–1012.

Stock, J. H., and M. W. Watson. 2005. "Implications of Dynamic Factor Models for VAR Analysis." NBER Working Paper No. 11467.

———. 2016. "Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics." In *Handbook of Macroeconomics,* Vol. 2A, ed. J. B. Taylor and H. Uhlig, 415–525 (chapter 8). Elsevier.

Monetary Policy, Commodity Prices, and Misdiagnosis Risk*

Andrew J. Filardo,[a] Marco J. Lombardi,[a] Carlos Montoro,[b] and Massimo Minesso Ferrari[c]
[a]Bank for International Settlements
[b]Central Reserve Bank of Peru and Professor of CENTRUM Católica, Pontificia Universitad Católica del Perú
[c]European Central Bank

How should monetary policy respond to commodity prices when the underlying drivers are difficult to diagnose accurately? If monetary authorities misdiagnose commodity price swings as being driven primarily by external supply shocks when they are in fact driven by global demand shocks, the conventional wisdom—to look through the first-round effects of commodity price fluctuations—may no longer be sound policy advice.

To analyze this question, we employ the multi-economy DSGE model of Nakov and Pescatori (2010), which splits the global economy into commodity-exporting and non-commodity-exporting economies. In an otherwise conventional DSGE setup, commodity prices are modeled as changing endogenously with global supply and demand developments, including global monetary policy conditions. This framework allows us to explore the implications of monetary policy decisions when there is a risk of misdiagnosing the drivers of commodity prices.

*We are grateful to the editor and a referee for their helpful and constructive comments. We thank conference participants at the Central Bank of Brazil's XVII Annual Inflation Targeting Seminar, CEMLA's XX Annual Meeting, the 2017 Central Bank Research Association (CEBRA) Annual Meeting, the 2019 Annual International Journal of Central Banking Research Conference, and the Second Macroeconomic Modelling and Model Comparison Network at Stanford University. We also thank seminar participants at the Bank for International Settlements, the Bank of Canada, and the Bank of Japan for their comments. The views expressed here are those of the authors and do not necessarily reflect those of the Bank for International Settlements, the Central Reserve Bank of Peru, or the European Central Bank.

We first confirm that monetary authorities deliver better economic performance when they are able to accurately identify the global nature of the shocks, i.e., global supply and demand shocks, driving commodity prices. Moreover, we show that when it is difficult to identify these shocks, monetary authorities can minimize some of the adverse feedbacks from misdiagnoses by targeting core inflation. Finally, we highlight the implications of misdiagnosis risk in the case where the monetary authority misinterprets supply-driven increases in commodity prices as demand driven; the contraction in both output and core inflation is larger than in the case of an accurate diagnosis.

In light of recent empirical studies documenting the significance of global demand in driving commodity prices, these findings call for giving greater prominence to global factors in domestic monetary policymaking and highlight potential gains from focusing on accurate diagnoses of domestic and global sources of shocks.

JEL Codes: E52, E61.

1. Introduction

Over the past decade, global commodity prices have experienced wide swings, reaching historically high levels in the run-up to the Great Recession before plummeting as the global economy collapsed. Prices subsequently rebounded with the global economic recovery but then fell again amid significant policy concerns. While challenging, this type of volatility is not a new environment for policymakers. Even though most commodity prices remained broadly stable during the so-called Great Moderation, they were quite volatile in the 1970s amid geopolitical tensions that pushed oil price volatility to then-unprecedented levels.

It was the experience of the 1970s that forged the conventional wisdom about how monetary authorities should respond to commodity price fluctuations. Commodity price fluctuations were seen largely as the result of exogenous supply shocks; in such an environment, the conventional wisdom that emerged was that, when facing such swings, monetary authorities should look through the first-round price effects and only respond to the second-round effects on wage and inflation expectations. In practice, this suggested a monetary policy focus on core inflation.

Views about the drivers of global commodity price swings have been evolving, especially in recent years, as a growing body of statistical evidence points to a new interpretation of commodity price swings. Kilian (2009), for example, finds evidence that oil price fluctuations have been increasingly influenced by demand from commodity-hungry emerging market economies (EMEs). The most recent literature has not challenged this view: Kilian and Baumeister (2016) argue that the oil price decline in 2015 should also be ascribed to a slowdown in global economic activity; and Stuermer (2017) and Fueki et al. (2018) emphasize the role of demand shocks in a historical perspective. In a similar vein, Sussman and Zohar (2016) report that commodity price fluctuations can be taken as a proxy for global demand. In a broader context, Filardo and Lombardi (2014) note the growing prominence of these global demand shifters for EME inflation dynamics. This evidence raises doubts about the relevance of the conventional wisdom.

The prominence of *endogenous* commodity price swings has important implications for monetary policy, given its central role in influencing aggregate demand. The relationship between monetary policy decisions and endogenous commodity prices implies an important two-way link. Monetary policy decisions influence aggregate demand and hence commodity prices. Indeed, Anzuini, Lombardi, and Pagano (2013) and Filardo and Lombardi (2014) report evidence that loose monetary policy has had an effect on commodity prices via the global demand channel.[1] At the same time, commodity price swings influence price stability and hence monetary policy decisions.

Given the global nature of commodity prices, misdiagnosis risk is particularly high in a world of many central banks with purely domestic monetary policy mandates. Individual countries may think that because they are sufficiently small, they can reasonably ignore the effect of their own policy decisions on the rest of the world,

[1]There is evidence that U.S. monetary policy plays a special role. Akram (2009) finds that lower interest rates in the United States boost commodity prices via an exchange rate channel.

and hence on commodity prices. This would be the case if all economies were hit by uncorrelated idiosyncratic shocks. However, global shocks imply that central banks are likely to respond in a correlated way which would endogenously feed back on commodity prices. A failure to internalize the endogenous feedbacks could contribute to economic and financial stability concerns (Caruana, Filardo, and Hofmann 2014; Rajan 2015).

From a modeling perspective, this discussion suggests the importance of developing monetary policy models with endogenously determined commodity prices and monetary authorities that are subject to misdiagnosis risk. To date, the bulk of the theoretical literature has stayed clear of models with endogenous commodity prices (see, e.g., Leduc and Sill 2004, Carlstrom and Fuerst 2006, Montoro 2012, Natal 2012, and Catao and Chang 2015). Moreover, this literature has generally focused on how a monetary authority should respond to exogenous movements in oil prices, e.g., whether it is optimal to target core or headline inflation and whether commodity price movements have far-reaching implications for the tradeoff between stabilizing output and controlling inflation. For example, Blanchard and Galí (2010) have gone so far as to argue that an increase in commodity prices driven by foreign demand can still be treated by a domestic monetary authority as an external supply shock. Such a conclusion is less tenable in models of endogenous commodity prices and correlated monetary policy reaction functions.

Various theoretical papers have addressed the endogeneity of commodity prices in small-scale dynamic stochastic general equilibrium (DSGE) models (e.g., Backus and Crucini 2000; Bodenstein, Erceg, and Guerrieri 2008; and Nakov and Nuño 2013). However, these models have generally ignored monetary policy, focusing instead on oil price determination and the frictions affecting it. Nakov and Pescatori (2010) is an early attempt to characterize monetary policy tradeoffs in a DSGE model in which oil prices are determined endogenously. Another important contribution to this literature is Bodenstein, Guerrieri, and Kilian (2012), who highlight, as we do, the benefits of identifying the nature of the shocks hitting the economy.

Our model extends this class of models by considering the policy challenges facing a monetary authority when it tries to infer the source of commodity price shocks.[2] Namely, there is a risk that a monetary authority may misdiagnose a commodity price swing as being driven by an external supply shock when it is, in fact, driven by an endogenous global demand shock, and vice versa. In our model, the commodity price is endogenously determined in equilibrium by the interplay of global demand and commodity supply from two types of commodity-exporting economies—one competitive and one monopolistic. In this setting, the optimal monetary policy response to commodity price swings depends on the perception of the underlying drivers of the swings. Unable to fully know the nature of the drivers, the monetary authority infers them via signal extraction, thereby opening up the possibility of systematic misdiagnoses.

The modeling exercise delivers several policy-relevant implications. First, it is important to distinguish between global demand and supply shocks when responding to commodity prices. If it is possible to accurately diagnose the source of a shock, our model finds that the best response to demand shocks is to lean against them fully (a result consistent with a standard New Keynesian closed-economy model). The best response to commodity supply shocks (i.e., a decrease in commodity prices) is to look through them.

Second, the conventional wisdom of looking through the first-round effects of commodity price swings is not always optimal. In our model, this result arises because our model breaks the "divine coincidence" between inflation and output gap stabilization (e.g., Blanchard and Galí 2007), which is a standard feature of DSGE models with exogenous commodity prices. The breaking of the divine coincidence comes, in part, from the assumption of a monopolistically competitive commodity exporter, and in part from the imperfect information environment.

Third, misdiagnosis risk matters. In the case where the monetary authority misinterprets supply-driven increases in commodity prices as demand driven, the contraction in both output and core inflation is larger than in the case of an accurate diagnosis. This indicates another reason for the breakdown of the divine coincidence in this model (even if the dominant exporter acts as a price taker). This

[2]See Filardo and Lombardi (2014) for a discussion of commodity price misdiagnosis risks in the context of Asian EMEs.

result underscores the potential benefits of trying to correctly diagnose the sources of commodity price swings when setting monetary policy. For example, a monetary authority that misdiagnoses global demand shocks as external supply shocks amplifies cyclical fluctuations (including commodity prices) and, as a result, destabilizes the economy.

2. Outline of the Baseline Model

We present a global monetary policy model in which commodity prices are determined endogenously, in the spirit of Nakov and Pescatori (2010, hereafter NP). The global economy is split into commodity-importing countries and commodity-exporting ones.

The commodity-importing countries are treated as one representative economy. It imports the commodity both for consumption and as an input in production of final goods and services. The firms produce final goods and services in a monopolistically competitive way in the face of nominal rigidities.

The commodity-exporting economies comprise a dominant commodity-exporting economy and a fringe of smaller competitive exporters. The dominant commodity-exporting economy has market power and sets prices above marginal cost. The fringe of small exporting countries operates competitively, taking the global commodity price as given. Consumers in these commodity-exporting countries buy final goods and services from the commodity-importing economy.[3]

The role of the central bank is to set monetary policy à la Taylor (1993). We extend the NP model by considering two types of uncertainty that the central bank faces. The first is real-time data uncertainty. Because inflation and output are typically only observed with a lag, we model the central bank as reacting to *expected* inflation and the *expected* output gap. The central bank can observe and

[3]Note that cross-border financial autarky is assumed, implying that current accounts are balanced in each period. Also, trade is assumed to be carried out in a common global currency, suppressing potential tradeoffs from exchange rate dynamics. These assumptions streamline the analysis and allow us to highlight the key implications of misdiagnosis risk which would be more complex in a richer model.

respond to commodity prices in real time. Relaxing the perfect information assumption opens up one way, within this class of models, for us to explore the consequences of central bank miscues in setting the policy rate. Second, and much less frequently addressed in the literature, is uncertainty about whether the commodity price is driven by demand or supply shocks. We assume that central banks cannot directly observe the source of commodity price shocks, but policymakers use available data to infer them. We explore the inferential challenges and implications that arise from the misdiagnosis of shocks, thereby shedding further light on the practical use of DSGE models.

The rest of this section sketches out the structure of the model. In addition to the two key points mentioned above, we also deviate from the Nakov and Pescatori (2010) setup in three other respects: (i) we introduce the commodity good into the households' utility function, which allows us to consider nontrivial policy responses to differences between headline and core inflation; (ii) we interpret the commodity as a broad basket of commodities rather than focusing narrowly on oil;[4] and (iii) we solve the Nash (instead of the Ramsey) problem for the dominant producer so as to reflect realistic information constraints facing producers.[5]

2.1 Representative Households and Firms in the Commodity-Importing Economy

Imported commodities play two roles in this economy. For households, commodities enter the consumption basket; as mentioned above, we make this dependence explicit to highlight the monetary policy implications of headline and core inflation. For firms in the commodity-importing economy, commodities are used in the production of final goods and services.

---

[4]This is consistent with the view that commodity markets as a whole remain an important driver of short-run inflation dynamics and often reflect cartel-like behavior for metals (and minerals) and anti-competitive behavior for foodstuffs (see Organisation for Economic Co-operation and Development 2012).

[5]The computer code for the model is available at http://www.macromodelbase.com/.

2.1.1 Representative Households

The representative household has a utility function over consumption (C_t) and labor (L_t),

$$U_{t_0} = E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} exp\left(e_t\right) \left[\ln\left(C_t\right) - \frac{L_t^{1+v}}{1+v}\right], \tag{1}$$

where e_t is a preference shock v and is the inverse of the labor supply elasticity. Consumption is defined as a Cobb-Douglas aggregator of *final* goods and services, $C_{Y,\,t}$, and the commodity, $\mathfrak{M}_{C,t}$:

$$C_t = \left(C_{Y,t}\right)^{1-\gamma} \left(\mathfrak{M}_{C,t}\right)^{\gamma}. \tag{2}$$

$C_{Y,\,t}$ is a Dixit-Stiglitz aggregate of a continuum of differentiated goods and services, $C_{Y,\,t}(z)$, of the form $C_{Y,t} = \left[\int_0^1 C_{Y,t}\left(z\right)^{\frac{\varepsilon-1}{\varepsilon}} dz\right]^{\frac{\varepsilon}{\varepsilon-1}}$.

The household maximizes its intertemporal utility subject to the period budget constraint,

$$C_t = \frac{W_t\, L_t}{P_t} + \frac{B_{t-1}}{P_t} - \frac{1}{R_t}\frac{B_t}{P_t} + \frac{\Gamma_t}{P_t} + \frac{T_t}{P_t}, \tag{3}$$

where W_t is the nominal wage, P_t the price of the consumption good, B_t the end-of-period nominal bond holdings, R_t the riskless nominal gross interest rate, Γ_t the share of the representative household's nominal profits, and T_t the net transfers from the government.

2.1.2 Firms

Monopolistically competitive firms in the commodity-importing economy produce final goods and services using a Cobb-Douglas technology:

$$Y_t\left(z\right) = A_t L_t\left(z\right)^{1-\alpha} \mathfrak{M}_{Y,t}\left(z\right)^{\alpha}, \tag{4}$$

where $\mathfrak{M}_{Y,t}$ is the commodity input and α denotes the commodity share in production. The firms' cost-minimization problem implies an expression for the real marginal cost, given the real commodity market price, $Q_t \equiv P_{\mathfrak{M},t}/P_t$:

$$MC_t\left(z\right) = \left(\frac{W_t}{P_t}\right)^{1-\alpha} Q_t^{\alpha} \Big/ \left[A_t\left(1-\alpha\right)^{1-\alpha}\alpha^{\alpha}\right].^{6} \tag{5}$$

These first-order conditions imply the demand equations for the commodity:

$$\mathfrak{M}_{Y,t}\left(z\right) = \alpha\frac{MC_t}{Q_t}Y_t\left(z\right). \tag{6}$$

Under Calvo pricing, and following Benigno and Woodford (2005), the implied nonlinear Phillips curve for final goods and services—i.e., core inflation ($\Pi_{Y,t}$)—is

$$\theta\Pi_{Y,t} = 1-\left(1-\theta\right)\left(\frac{N_t}{D_t}\right)^{1-\varepsilon}, \tag{7}$$

where N_t/D_t is the equilibrium relative price, with $D_t = Y_t/C_t + \theta\beta E_t\left[\left(\Pi_{Y,t+1}\right)^{\varepsilon-1}\mathrm{D}_{t-1}\right]$ and $N_t = \mu Y_t MC_t/C_t + \theta\beta E_t\left[\left(\Pi_{Y,t+1}\right)^{\varepsilon}\mathrm{N}_{t-1}\right]$.

2.1.3 Aggregation

Aggregating firms' demand for labor and the commodity yields a set of equations with which to solve the model. Even though labor and commodity demand differ across firms due to Calvo staggered pricing, the aggregate-demand equations resemble those of the individual firm's demand equations except for an adjustment capturing the effect of price dispersion. Higher inflation increases the dispersion of prices, and this increased price dispersion boosts labor and commodity demands for a given level of output.

[6]Note that all firms face the same real marginal costs given the constant-returns-to-scale technology and competitive factor markets. In this model, the real commodity price is proportional to the inverse of the importing economy's terms of trade.

2.2 Representative Households and Firms in the Commodity-Exporting Economies

The commodity market is modeled as comprising a dominant commodity exporter and a fringe group of competitive commodity exporters. The dominant exporter produces monopolistically, and the fringe produces competitively. For completeness, the households in these economies face conventional decision problems.

2.2.1 Dominant Commodity Exporter

In each period, the dominant exporter chooses its supply to maximize profits, taking global demand as given and internalizing the response of the competitive fringe. The dominant exporting economy produces the commodity according to the technology:

$$\mathfrak{M}_t = Z_t I_t^{*,D}, \tag{8}$$

where Z_t is an exogenous productivity shifter and $I_t^{*,D}$ is the intermediate input (bought from the commodity-importing economy). Productivity evolves exogenously according to

$$\ln Z_t = (1 - \rho_z) \ln \bar{Z} + \rho_z \ln Z_{t-1} + \varepsilon_t^z, \tag{9}$$

where $\varepsilon_t^z \sim i.i.d.N\left(0, \sigma_z^2\right)$. Shocks to Z_t can then be interpreted as *global commodity supply shocks.*

The household utility function depends only on consumption of final goods and services:

$$U_{t_0}^{*,D} = E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \ln\left(C_t^{*,D}\right) \tag{10}$$

subject to the period budget constraint, $P_{Y,t} C_t^{*,D} = \Gamma_t^{*,D}$, which equates consumption expenditure to dividends from commodity production, $\Gamma_t^{*,D}$.

2.2.2 Fringe of Competitive Commodity Exporters

The fringe exporters comprise a continuum of atomistic firms indexed by $j \in [0, \Omega_t]$. Each produces a quantity $X_t(j)$ of the commodity according to technology of the form

$$X_t\left(j\right) = \xi\left(j\right) Z_t I_t^{*,F}\left(j\right), \text{ subject to capacity constraint}$$
$$X_t\left(j\right) \in \left[0, \bar{X}\right], \tag{11}$$

where $\left[\xi\left(j\right) Z_t\right]^{-1}$ is the marginal cost of economy j defined by an idiosyncratic shock and global supply shock. Input $I_t^{*,F}\left(j\right)$ is an intermediate input used in commodity production and is bought from the commodity-importing economy. Each firm maximizes profits, taking as given the international real price of the commodity, Q_t,

$$\max Q_t X_t\left(j\right) - \frac{\frac{P_{Y,t}}{P_t} X_t\left(j\right)}{\xi\left(j\right) Z_t} s.t.\, X_t\left(j\right) \in \left[0, \bar{X}\right]. \tag{12}$$

The resulting supply by the fringe of competitive exporters is $X_t = \Omega_t Z_t Q_t$.

Like the setup for the dominant exporting economy, the household utility function in the fringe economies depends only on the aggregate consumption of final goods and services:

$$U_{t_0}^{*,F} = E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \ln\left(C_t^{*,F}\right), \tag{13}$$

subject to the following period budget constraint, $P_{Y,t} C_t^{*,F} = \Gamma_t^{*,F}$, which equates consumption expenditures to dividends earned by the fringe commodity exporters, $\Gamma_t^{*,F}$.

2.2.3 Aggregate Production of the Commodity

In the case of a perfectly competitive fringe market of commodity exporters, the equilibrium commodity price is equal to marginal costs:

$$Q_t^{PC} = Z_t^{-1}, \tag{14}$$

and the quantity produced is given by the global demand at that price. Having market power, the dominant commodity exporter determines its supply by internalizing the optimal response of the fringe and the global demand function for the commodity. The dominant producer's objective function is

$$\max_{\{\mathfrak{M}_t\}_{t_0}^{\infty}} E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \ln\left(Q_t^{1/1-\gamma}\mathfrak{M}_t - \mathfrak{M}_t/Z_t\right) \tag{15}$$

given commodity demand and supply from fringe suppliers, $\mathfrak{M}_t = \mathfrak{M}_{C,t} + \mathfrak{M}_{Y,t} - X_t$.[7]

The first-order condition of this problem determines the commodity price (in real terms):

$$Q_t = \Psi_t Z_t^{-1}, \tag{16}$$

where $\Psi_t \equiv 1/(1-\eta_t)$ is the commodity market markup and $\eta_t = \mathfrak{M}_t/(\mathfrak{M}_t + 2X_t)$ is the elasticity of substitution of the global demand for the commodity (in absolute value). Accordingly, the commodity price is a markup over marginal cost, the latter being influenced by commodity supply shocks, firm-specific supply shocks, and shifts in g_t.[8]

2.2.4 Market Clearing

Under the assumption that all relevant markets clear, there is a well-defined aggregate demand for final goods and services equal to aggregate supply,

$$Y_t = C_{Y,t} + C_t^{*,D} + I_t^{*,D} + +C_t^{*,F} + +I_t^{*,F} + g_t, \tag{17}$$

[7]Note that the dominant exporter takes as given the macroeconomic variables of the fringe producers, e.g., $(C, MC, Y, \Omega, \Delta)$, and internalizes the fringe's reaction, but not the feedbacks of its actions on the macroeconomic performance of the commodity-importing economy. Nakov and Pescatori (2010) analyze this case in which the dominant exporter completely internalizes its actions on the importing country. Technically, they solve the Ramsey problem of the dominant exporter, which may be useful in capturing longer-run supply behavior in the data. Our Nash solution is more restrictive but may be seen as more reasonable when considering a central bank's short-run tradeoffs in a world with incomplete information. Further details on the dominant exporter's problem are provided in appendix B.

[8]In addition, the markup $\Psi_t \equiv 1/(1-\eta_t)$ is an increasing function of the dominant commodity exporter's market share relative to that of the competitive fringe of commodity exporters. The limiting case is when $\mathfrak{M}_t \rightarrow 0$ corresponds to perfect competition while $X_t \rightarrow 0$ is the case of a single monopolist.

which includes final goods consumption in the commodity-importing economy and the aggregate consumption and intermediate goods demanded by the dominant and the competitive fringe of the commodity-exporting countries, respectively (superscript D denotes the dominant commodity-exporting economy and F the competitive fringe).[9] Finally, g_t[10] is an aggregate demand shock, which captures a positive shift in the demand for the final good produced by the commodity-importing country. Note that the monetary authority cannot fully offset the effect of this shock in the same way as it could with a markup shock.

2.3 Characterizing Monetary Policy

Monetary policy in the commodity-importing country is modeled as a linear Taylor-type rule informed by deviations from model-consistent benchmarks for output, inflation, and the interest rate. This section defines the benchmarks and analyzes the implications of alternative policy rules.

2.3.1 Optimal Benchmarks

Benchmark output gaps can be derived by substituting the equations for labor demand, labor supply, aggregate demand, and commodity demand for production into the aggregate production function. The log-level of output in terms of marginal costs, the dispersion of prices, productivity, and the real commodity price is[11]

$$y_t = \frac{1}{1-\alpha} a_t + \frac{\alpha}{1-\alpha}(mc_t - q_t) + \frac{1}{1+v}\Upsilon\left(mc_t + \frac{\gamma}{1-\gamma} q_t\right). \tag{18}$$

[9]The structural parameters in the model are calibrated to be in line with those found in the literature (see table A.1 in appendix A). Fiscal policy is modeled as a linear rule: $T_t = \tau P_{Y,t} Y_t$.

[10]g_t is modeled as an AR(1) process with zero mean, i.e., $g_t = \rho_g g_{t-1} + \varepsilon_t^g$.

[11]The level of output is $Y_t = \left(\frac{A_t}{\Delta_t}\right)^{1/(1-\alpha)} \left[\frac{(1-\alpha) MC_t \Delta_t}{(Q_t)^{-\gamma/(1-\gamma)} - \alpha MC_t \Delta_t}\right]^{1/(1+v)} \left[\alpha \frac{MC_t \Delta_t}{Q_t}\right]^{\alpha/(1-\alpha)}$ and $\Upsilon \equiv \left[1 - \frac{\alpha}{\mu} Q^{\gamma/(1-\gamma)}\right]^{-1} \geq 1$. A full derivation is provided in appendix C.

The *level of natural output*, y_t^n, is defined as the level of output consistent with a flexible-price equilibrium. In this case, the marginal cost is constant, $MC_t = \mu^{-1}$, and there is an absence of price dispersion, $\Delta_t = 1$. In log-linear terms, the level of natural output, y_t^n, is of the form

$$y_t^n = \frac{\alpha}{1-\alpha} a_t - \left(\frac{\alpha}{1-\alpha} - \frac{1}{1+v} \frac{\gamma}{1-\gamma} \Upsilon \right) q_t. \tag{19}$$

As shown in equation (19), commodity price fluctuations have two opposing effects on the level of natural output. From the side of production (i.e., the first term in the parentheses), an increase in the commodity price has a qualitative effect similar to that of a negative productivity shock; it reduces the level of natural output. From the side of consumption (i.e., the second term in the parentheses), an increase in the commodity price increases the level of natural output. The latter term reflects an increase in labor due to a negative income effect from a higher commodity price.

The *natural output gap*, $\hat{y}_t^n = \left[\frac{\alpha}{1-\alpha} - \frac{1}{1-v} \Upsilon \right] mc_t$, measures the difference between the actual and the natural level of output. This implies that responding to the natural output gap is equivalent to responding to real marginal costs, up to a scale factor.

Similarly, the *log-level of efficient output*, y_t^e, is defined with respect to the efficient allocation, i.e., flexible prices and no monopolistic distortions in the commodity market or in the final goods market (which implies that $Q_t^e = Z_t^{-1}$ and $\mu^e = 1$):[12]

$$y_t^e = \frac{\alpha}{1-\alpha} a_t - \left(\frac{\alpha}{1-\alpha} - \frac{1}{1+v} \frac{\gamma}{1-\gamma} \Upsilon^e \right) z_t. \tag{20}$$

A key difference is that commodity markup shocks do not affect the level of efficient output: such output is instead affected only by fluctuations associated with supply shocks in the commodity market. As a consequence, a demand-driven increase in the commodity price

[12]With $\Upsilon^e \equiv \left[1 - \alpha Z^{-\gamma/(1-\gamma)} \right]^{-1}$. The relationship between Υ^e and Υ depends on the extent of monopolistic distortions. Υ^e and Υ are equal only if both markets are perfectly competitive or if the commodity is not used for production (that is, $\alpha = 0$).

would leave the benchmark efficient output gap unchanged. However, a negative commodity supply shock would decrease both the natural and efficient output levels, albeit by different amounts.[13] The efficient output gap, $\hat{y}_t^e$, which is defined as the difference between actual output and the efficient level of output, is of the form[14]

$$\hat{y}_t^e = \hat{y}_t^n - \left(\frac{\alpha}{1-\alpha} - \frac{1}{1+v}\frac{\gamma}{1-\gamma}\Upsilon\right)\hat{\Psi}_t - \frac{1}{1+v}\frac{\gamma}{1-\gamma}(\Upsilon - \Upsilon^e)Z_t, \tag{21}$$

where this is the welfare-relevant output gap and is equal to the natural output gap plus a term that depends on the commodity price markup and the commodity supply shock.

2.3.2 Breaking Down the Divine Coincidence

Both core inflation and headline inflation are determined by the natural output gap, expected inflation, and commodity price changes. Expressed in log-linear terms, the equations for core inflation and headline inflation are, respectively,

$$\pi_{Y,t} = \kappa_y \hat{y}_t^n + E_t \pi_{Y,t+1} \text{ and} \tag{22}$$

$$\pi_t = \pi_{Y,t} + \frac{\gamma}{1-\gamma}\Delta q_t. \tag{23}$$

Equation (22) describes the determinants of core inflation, and equation (23) describes headline inflation written in the form of a Phillips curve for aggregate final goods with $\hat{y}_t^n$ the natural output gap. Stabilization of the natural output gap is equivalent to stabilization of core inflation. And, in that case, headline inflation would vary proportionally with changes in real commodity prices. Equation (22) can be written in terms of the efficient output gap:

$$\pi_{Y,t} = \kappa_y \hat{y}_t^e + E_t \pi_{Y,t+1} + u_t, \tag{24}$$

[13]The level of efficient output contracts less than the level of natural output in response to a negative commodity supply shock because the commodity price markup partially offsets the effects of supply shocks on the commodity price.

[14]$\hat{\Psi}_t$ and z_t are Ψ_t and Z_t expressed in log-deviations from the steady state; $\Upsilon^e \equiv \left[1 - \alpha Z^{-\gamma/(1-\gamma)}\right]^{-1}$.

where u_t is an endogenous cost-push shock, which is a function of both $\hat{\Psi}_t$ and z_t. In this model, the divine coincidence featured in models with exogenous commodity prices is broken. It is no longer possible to simultaneously stabilize core inflation and the welfare-relevant output gap. The tradeoff arises from the effect of commodity price fluctuations on the level of efficient output. An increase in commodity price markups generates a positive cost-push shock, which puts upward pressure on core inflation but lowers the efficient output gap.

3. Monetary Policy Responses to Commodity Prices

With the model outlined in section 2 (and calibrated as reported in appendix A), we now explore the performance of alternative monetary policy rules assuming different types of imperfect information. This allows us to focus on both the possibility that central banks may misdiagnose the nature of the shocks hitting an economy and the associated implications. The imperfect information settings also highlight the absence of the divine-coincidence property: monetary policy cannot perfectly stabilize inflation and the output gap unless the central bank is able to accurately identify the nature of underlying shocks.[15]

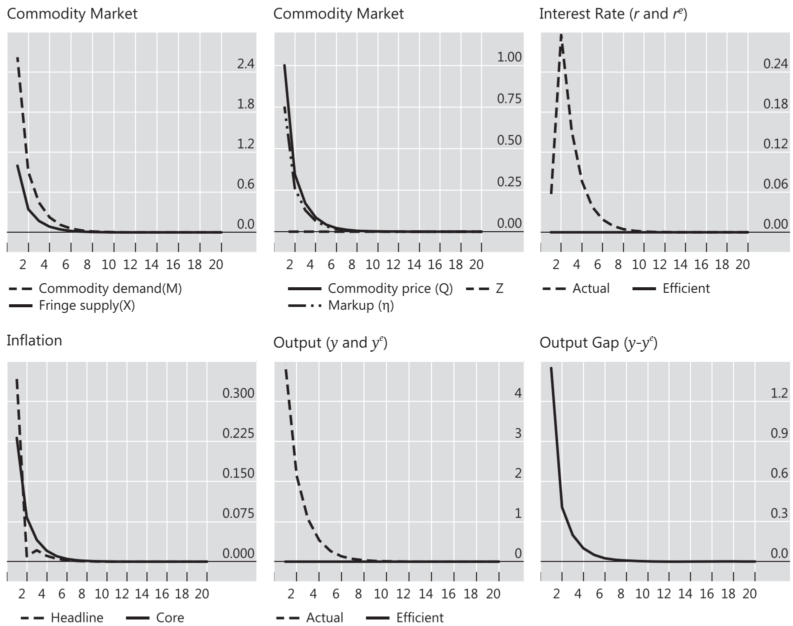
3.1 Monetary Policy Responses under Conventional Data Uncertainty

The baseline policy rule assumes that the monetary authority responds to expectations of core inflation and the efficient output gap. It is a log-linear Taylor-type rule:

$$r_t = E_{t|t-1}\left[r_t^e + \varphi_{core}\pi_{Y,t} + \varphi_y\left(y_t - y_t^e\right)\right] + \varphi_{com}\Delta q_t, \tag{25}$$

[15] Consistent with the Nakov and Pescatori (2010) model, we focus on the implications of uncertainty for monetary policy tradeoffs only in the commodity-importing economy by assuming nominal rigidities are absent in the commodity-exporting countries.

Figure 1. Responses to a Positive Aggregate-Demand Shock (ε_t^g)



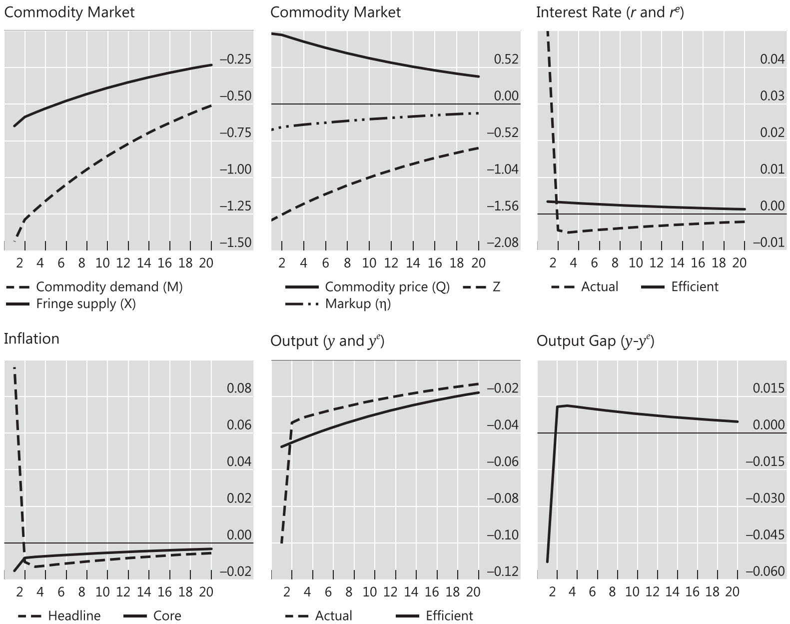
Notes: Impulse responses to an aggregate-demand shock using policy rule $r_t = E_{t|t-1}(r_t^e + \varphi_{core}\pi_{Y,t} + \varphi_y \hat{y}_t^e) + \varphi_{com}\Delta q_t$. The shock is calibrated so as to generate a 1 percent increase in commodity prices.

which includes a benchmark interest rate (r_t^e), core inflation ($\pi_{Y,t}$), and output gap $\hat{y}_t^e = (y_t - y_t^e)$. We assume that the monetary authority responds to *expected* inflation and output gap, whereas the change in the commodity price (Δq_t) is available in real time.

The performance of policy rule (25) is graphically assessed using impulse responses to demand and supply shocks.[16] Figure 1 plots the responses to a positive aggregate demand shock, defined as an exogenous shift in aggregate demand through g_t in equation (17). Figure 2 plots the response to a negative supply shock, defined as

[16]The coefficients of the policy rule are set to $\varphi_{core} = 1.5$, $\varphi_y = 0.5$, and $\varphi_{com} = 0.05$.

Figure 2. Responses to a Negative Supply Shock (ε_t^z)



Notes: Impulse responses to a (negative) supply shock. Other details are in figure 1.

an exogenous shift in Z_t in equation (9). Both shocks are calibrated to produce, on impact, a 1 percent increase in the commodity price.

Following a demand shock (figure 1), the commodity price rises (top-center panel). The demand shock also drives up output (but not the level of the efficient output) and the output gap. Both core and headline inflation initially rise. With the efficient interest rate unchanged, the policy rule calls for an increase in the policy rate (top-right panel). For roughly six quarters, the interest rate and output gap remain elevated. Note that the intuitively plausible shapes of the response underscore the reasonableness of our calibration for policy analysis.

A negative supply shock (figure 2) drives up the commodity price, but the policy rate response as well as the macroeconomic ones are modest. This constellation of responses confirms that the

Table 1. Expected Welfare Loss for Alternative Efficient Policy Rules, Percent of Steady-State Consumption ($\times$ E – 4)

	Model				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Welfare Loss →	**–3.80**	**–6.05**	**–5.91**	**–4.98**	**–5.27**
Policy Rule Specifications for Each Model					
φ_{core}	2.0	—	1.5	1.5	2.0
φ_{head}	—	2.0	—	—	—
φ_y	—	—	0.5	0.5	1.0
φ_{com}	—	—	—	0.05	0.1

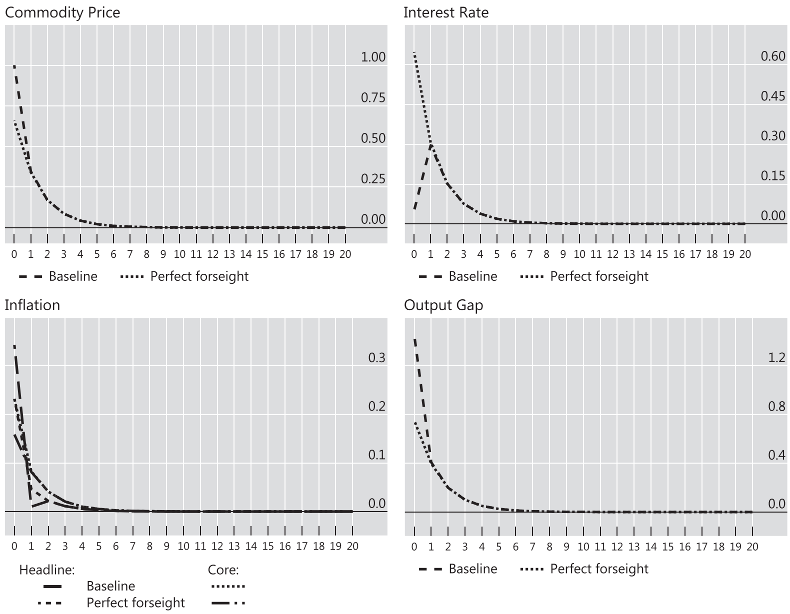
Notes: The unconditional expected welfare, EW_t, is assessed using a second-order solution of the model, where $W_t = U(C_t, L_t, e_t) + \beta E_t W_t$. The welfare cost in terms of steady-state consumption is equivalent to $\{exp((1-\beta)(EW_t - W))\} \times 100$. The Taylor-type monetary policy rule is specified as $r_t = E_{t|t-1}[r_t^e + \varphi_{core}\pi_{Y,t} + \varphi_{head}\pi_t + \varphi_y(y_t - y_t^e)] + \varphi_{com}\Delta q_t$.

baseline model (with its imperfect information assumption) delivers the conventional wisdom that monetary authorities should essentially look through the first-round effects of supply shocks.

Table 1 provides an alternative metric of model performance based on the model's welfare criterion under different policy rules. The first column corresponds to a rule with a response only to expected core inflation; it achieves the lowest (expected welfare) loss. Columns 2–5 correspond to alternative policy rules. The rule in column 2 considers a response only to expected headline inflation; this produces the worst outcome, indicating the superiority of targeting core versus headline inflation in our model. Column 3 introduces an additional response to the expected output gap, and columns 4 and 5 add a direct response to the commodity price change.[17]

[17]One drawback of evaluating expected welfare losses with simple linear rules is that the results may be sensitive to the variances of the shocks. We have conducted a battery of tests to confirm that our findings are robust to a range of reasonable variances.

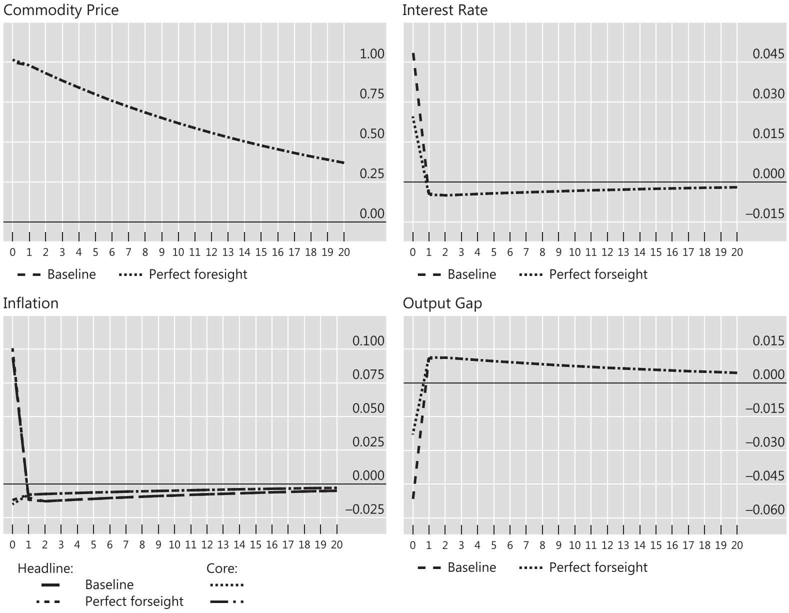
Figure 3. Response to a Positive Aggregate-Demand Shock under Full Information



Notes: Impulse responses to an aggregate-demand shock using the baseline policy rule of figure 1 and the full-information rule $r_t = r_t^e + \varphi_{core}\pi_{Y,t} + \varphi_y (y_t - y_t^e) + \varphi_{com}\Delta q_t$. The shock is calibrated so as to generate a 1 percent increase in commodity prices under the baseline policy rule.

While results suggest that responding to commodity prices can improve performance over responding to headline inflation, they leave open questions such as whether monetary authorities can achieve even better performance by inferring the nature of the shocks driving commodity price developments. Data uncertainty indeed induces substantial changes in the reaction to shocks. To illustrate this, we compare the baseline results with those under full information, i.e., when the monetary authority is able to observe with certainty macrovariables at each point in time. Figures 3 and 4 compare the reaction of the system to demand and shocks with the baseline. The overall picture is that first-round effects generated by uncertainty are already sizable. Looking at the aggregate-demand shock (figure 3) highlights that under full information the monetary

Figure 4. Response to a Negative Supply Shock under Full Information



Notes: Impulse responses to a supply shock using the baseline policy rule of figure 1 and the full-information rule $r_t = r_t^e + \varphi_{core}\pi_{Y,t} + \varphi_y\left(y_t - y_t^e\right) + \varphi_{com}\Delta q_t$. The shock is calibrated so as to generate a 1 percent increase in commodity prices under the baseline policy rule.

authority increases the interest rate more aggressively compared with the baseline. This is no surprise, given that the monetary authority is fully aware of the effects of the shock on output. By doing so, the monetary authority not only dampens cyclical fluctuations but also moderates the increase in commodity prices and hence in inflation. Also, under a supply shock (figure 4) full information helps the monetary authority stabilize output, although in this case differences are less striking.

3.2 Monetary Policy Responses under Misdiagnosis Risk

In addition to uncertainty about current output and inflation, when setting policy rates monetary authorities do not have precise knowledge of the underlying nature of the shocks affecting commodity prices. This section considers the challenges of inferring the nature of the shocks and implications arising from possible misdiagnoses, i.e., the risk that a monetary authority may misinterpret the source of the shock driving changes in commodity prices.

We cast the problem of shock identification in a classical signal-extraction problem.[18] Starting with the assumption that the monetary authority does not observe supply and demand shocks (ε_t^z and ε_t^g), it can infer them from available observations. Consider a simple linear model for commodity prices:

$$q_t = -z_t + \psi_t = H_t' \xi_t, \tag{26}$$

where z_t and ψ_t are the supply and the demand shock in the central banks' signal-extraction model, $H_t = [-1 \quad 1]$, and $\xi_t = [z_t \quad \psi_t]'$. The unconditional variance of ξ_t is

$$P \equiv var\left(\xi_t\right) = \begin{bmatrix} \sigma_z^2 & \sigma_{z\psi} \\ \sigma_{z\psi} & \sigma_\psi^2 \end{bmatrix}. \tag{27}$$

Given this informational structure, the monetary authority infers the sources of commodity price fluctuations by solving a signal-extraction problem using a Kalman filter, i.e.,

$$E_t^{ma}\left[z_t \quad \psi_t\right]' = Mq_t, \tag{28}$$

where $M = PH\left[H_t' PH\right]^{-1}$ is a weighted average of the variances and covariances of z_t and ψ_t; M is calculated as

$$M = \frac{x}{x^2 - 2\rho x + 1} \begin{bmatrix} \rho - x \\ \frac{1}{x} - \rho \end{bmatrix}, \tag{29}$$

where $\rho = corr\left(z_t, \psi_t\right)$ and $x = \sigma_\psi / \sigma_z$.[19]

Three cases of equation (28) shed light on the tradeoffs facing the monetary authority. In the first case (type A), when $x \to 0$, the volatility of the commodity supply shock is high relative to that of

[18] An online appendix, available at http://www.ijcb.org, also describes the Bayesian learning approach to the problem.

[19] See chapter 13 of Hamilton (1994) for the derivation.

the commodity market markup. In this case, the monetary authority attributes nearly all of the fluctuations to demand shocks. That is,

$$if\ x \to 0, \qquad E_t^{ma}\left[z_t \quad \psi_t\right]' \to \left[0 \quad q_t\right]'.$$

In the second case (type B), the fluctuations are attributed to supply shocks. That is,

$$if\ x \to \infty, \qquad E_t^{ma}\left[z_t \quad \psi_t\right]' \to \left[-q_t \quad 0\right]'.$$

In the last case (type C), the monetary authority attributes the commodity price fluctuations partially to each component of the commodity price, taking into account the relative volatility and correlation in equation (29).

Armed with these expectations, we can rewrite the monetary authority's policy rule as

$$\begin{aligned} r_t &= E_{t|t-1}^{ma}\left[r_t^e + \varphi_{core}\pi_{Y,t} + \varphi_y \hat{y}_t^e\right] + \varphi_{com}\Delta q_t \\ &= E_{t|t-1}\left[r_t^e + \varphi_{core}\pi_{Y,t} + \varphi_y \hat{y}_t^e\right] + \varphi_{com}\Delta q_t + e_t, \end{aligned} \tag{30}$$

where

$$\begin{aligned} e_t = &\left[E_{t|t-1}^{ma}\left(r_t^e\right) - E_{t|t-1}\left(r_t^e\right)\right] + \varphi_{core}\left[E_{t|t-1}^{ma}\left(\pi_{Y,t}\right) - E_{t|t-1}\left(\pi_{Y,t}\right)\right] \\ &+ \varphi_y\left[E_{t|t-1}^{ma}\left(\hat{y}_t^e\right) - E_{t|t-1}\left(\hat{y}_t^e\right)\right]. \end{aligned}$$

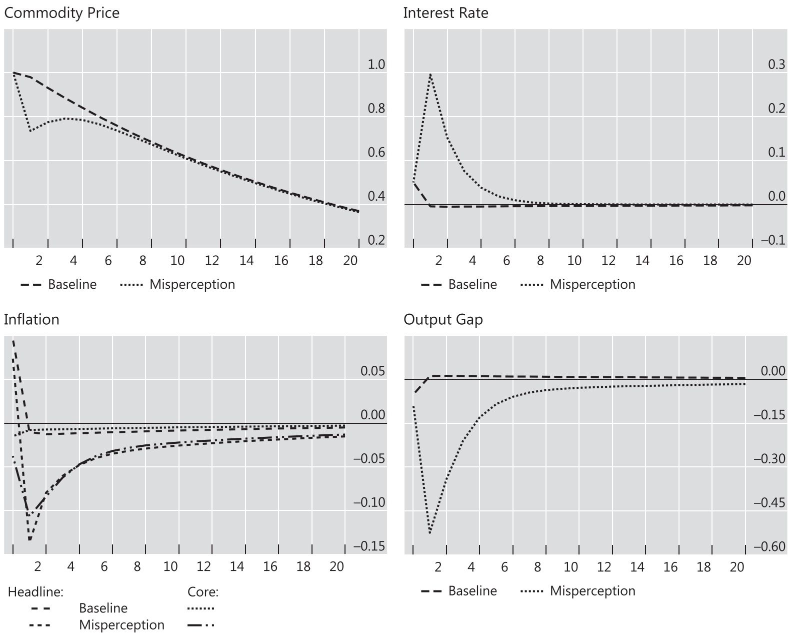
The e_t corresponds to a misdiagnosis error, which is *endogenous,* and $E_{t|t-1}^{ma}$ denotes the expectations under the incorrect diagnosis on the source of the shock. Note that when the monetary authority imputes the change in commodity prices to the wrong type of shock, its estimates of endogenous variables will be incorrect, leading to persistent errors in interest rate setting.[20]

To investigate the implications of the endogenous error, we replace equation (25) in our baseline model with equation (30). The resulting impulse responses highlight the findings.

Figure 5 shows the impulse responses to a commodity supply shock in the misdiagnosis case A. Even though the commodity price

[20] If the monetary authority correctly identifies the source of the shock, the error is zero.

Figure 5. Responses to a Negative Supply Shock under Misdiagnosis Type A

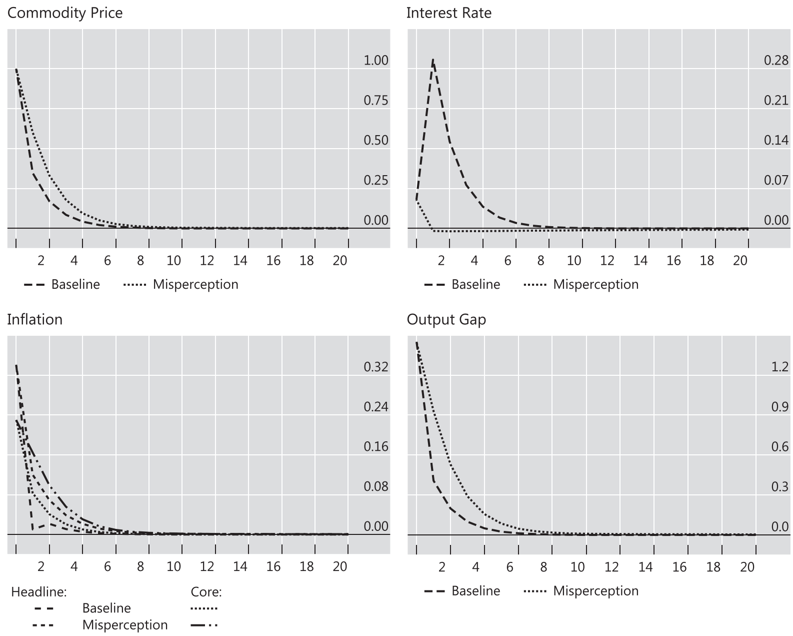


Notes: Impulse responses to a negative commodity supply shock when the monetary authority attributes all the fluctuation in the commodity price to an aggregate-demand shock (misdiagnosis type A). Other details are in figure 1.

is driven by a supply shock, the monetary authority misdiagnoses it as a traditional demand-driven commodity shock. If the monetary authority fails to recognize that an increase in the commodity price is driven by (external) supply conditions, the consequence is overly tight monetary policy accompanied by a sizable drop in both output and inflation. The commodity price responds in a more muted way than in the baseline case because of tighter monetary policy. Core and headline inflation both fall because of the slowing economy.

Figure 6 displays the impulse responses to a conventional aggregate demand shock in the case of misdiagnosis type B, i.e., when the rise in the commodity's price is mistakenly attributed to a negative commodity supply shock. In this case, the easier monetary policy associated with the looking-through-the-supply-shock strategy

Figure 6. Responses to a Positive Aggregate-Demand Shock under Misdiagnosis Type B

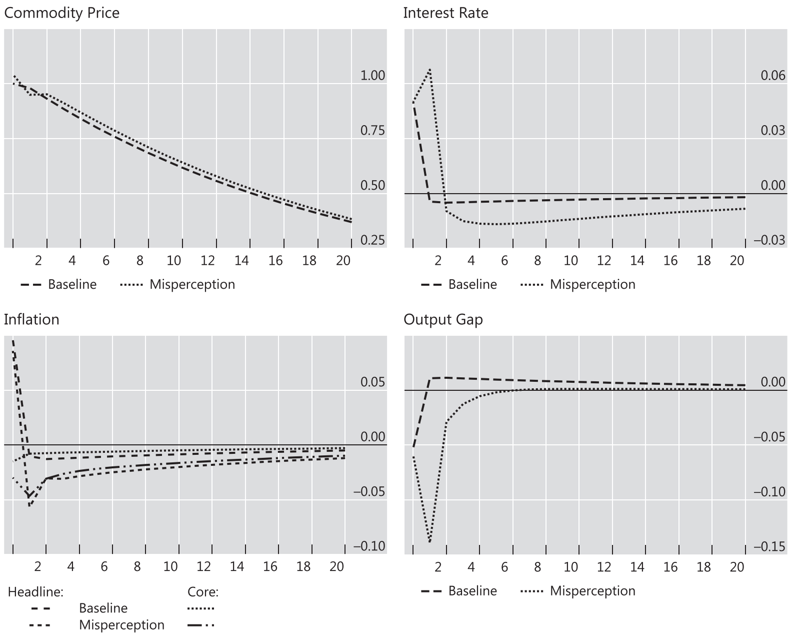


Notes: Impulse responses to an aggregate-demand shock when the monetary authority attributes all the fluctuation in the commodity price to a supply shock (misdiagnosis type B). Other details are in figure 1.

results in higher output and inflation. This type of policy misdiagnosis induces a procyclical increase in the commodity price.

Figures 7 and 8 report the results for the case of misdiagnosis type C, in which the monetary authority implements the optimally weighted response to the commodity price rise based on the historical commodity demand and supply shocks. The standard deviations of the shocks are calibrated to the empirical estimates by Filardo and Lombardi (2014), which yields a ratio of about 1.5 (x in equation (29)). Consistent with misdiagnosis cases A and B, the monetary authority responds excessively to supply shocks and insufficiently to demand shocks: on net, monetary policy appears excessively procyclical in the model.

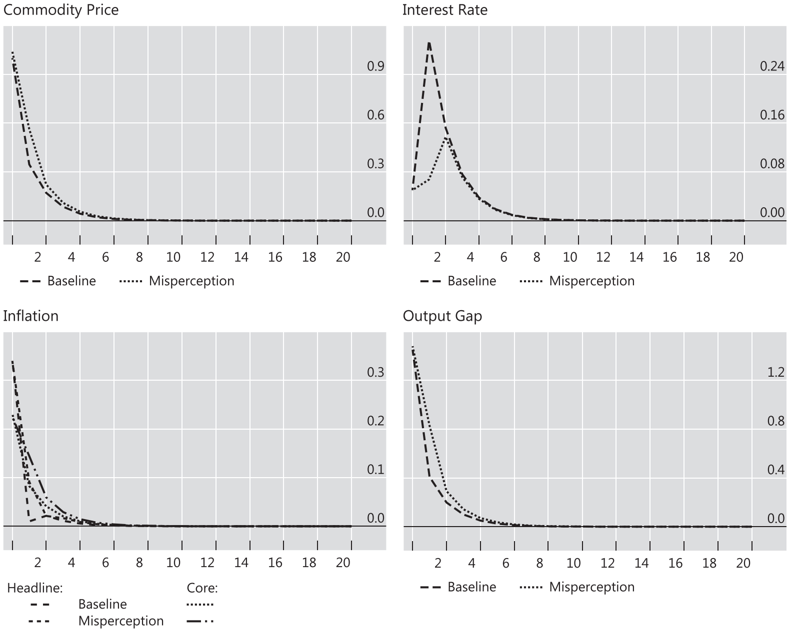
Figure 7. Responses to a Negative Supply Shock under Misdiagnosis Type C



Notes: Impulse responses to a negative commodity supply shock when the monetary authority attributes the fluctuation in the commodity price proportionally to aggregate-demand shock and commodity supply shock, with weights given by the ratio of their standard deviations (misdiagnosis type C). Other details are in figure 1.

The relative importance of these misdiagnosis risks can also be assessed by examining the welfare losses under different policy rules. The left-hand and right-hand panels of figure 9 plot, respectively, the welfare loss of the misdiagnosis types A and B cases (relative to the case of correct identification). In the case where the monetary authority misinterprets a rise in the commodity price as supply driven, the contraction in both output and core inflation is larger than in the full-information case. And, in the case where commodity price fluctuations are driven by global demand, the monetary authority amplifies cyclical fluctuations and, as a result, destabilizes the economy. Hence, it is no surprise that the expected welfare losses are always greater than zero under misdiagnoses.

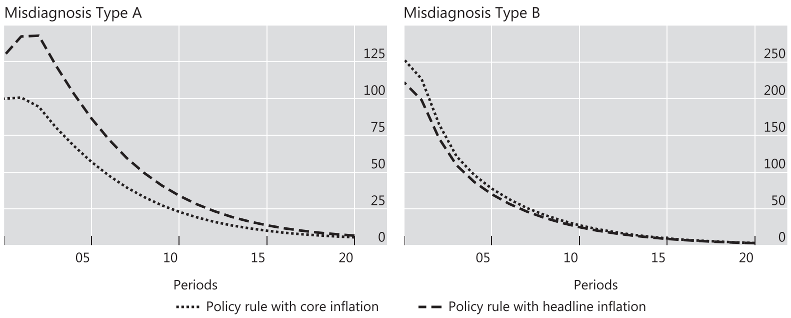
Figure 8. Responses to a Positive Aggregate-Demand Shock under Misdiagnosis Type C



Notes: Impulse responses to an aggregate-demand shock when the monetary authority attributes the fluctuation in the commodity price proportionally to aggregate-demand shock and commodity supply shock, with weights given by the ratio of their standard deviations (misdiagnosis type C), assuming the efficient benchmark policy rule and a 0.5 autoregression coefficient for the shock processes.

With respect to the performance of core and headline policy rules under these cases of misdiagnosis, the results still lend support to policy rules with core inflation rather than headline inflation. In misdiagnosis case A (supply shock treated as if it was demand), the core inflation rule implies a much lower expected loss than the headline one. This is not surprising given the more muted policy response associated with the headline inflation rule. In the misperception B case, the difference in the expected loss between the two rules is small, with a slight edge to the headline rule. If misdiagnosis risk cases A and B were both 50-50, the expected loss criteria

Figure 9. Ratio of the Welfare Losses under Misdiagnosis



Notes: Welfare is computed using a second-order solution of the model, where $W_t = U(C_t, L_t, e_t) + \beta E_t W_t$ and the welfare cost in terms of steady-state consumption is equivalent to $\{exp\,[(1-\beta)(EW_t - W) - 1]\} \times 100$.

would support the core inflation rule in this imperfect information environment.

Overall, the signal-extraction results reinforce the earlier findings on the importance of correctly identifying the underlying nature of commodity price shocks. On the modeling side, the misdiagnosis risk and its inherent procyclicality lead to a breakdown of the divine coincidence found in full-information models à la Blanchard and Galí (2007). On the policy side, the case for a core inflation rule is strengthened.

4. Conclusions

In this paper we documented that (i) monetary authorities deliver better economic performance when they are able to accurately identify the global nature of the shocks, i.e., global supply and demand shocks, driving commodity prices; and (ii) when it is difficult to identify these shocks, monetary authorities can minimize some of the adverse feedbacks from misdiagnosis by targeting core inflation. Global shocks also imply that central banks may find themselves responding in a correlated way. If so, it is important to account for this possibility. This reinforces arguments for greater prominence to be given to global factors in domestic monetary policymaking.

One important aspect of monetary policy misdiagnosis risk that deserves further research includes the important issue of parameter uncertainty, especially with respect to the slope of the Phillips curve. Indeed, recent empirical evidence suggests that the slope has flattened, and at least has become more uncertain, across many economies. In such a situation, the main call of this paper for more attention to be given to the nature of the shocks hitting the economy may take on even greater prominence. This may require going beyond the methods advocated in this paper, e.g., exploiting big data with cross-country coverage.

Appendix A. Model Calibration

Table A.1. Baseline Calibration

Structural Parameters	Parameter	Value
Share of Commodity in Consumption Basket	γ	0.10
Share of Commodity in Production Function	α	0.10
Inverse Frisch Labor Supply Elasticity	ν	0.50
Price Elasticity of Substitution	ε	7.66
Quarterly Discount Factor	β	0.99
Price Adjustment Probability	θ	0.75
Size of Competitive Commodity Production Relative to GDP	X/Y	0.10

Notes: For the parameters associated with the commodity market, the choice is a bit more arbitrary due to the less conventional view of the literature. The share of the commodity in the consumption basket is set to 10 percent, which roughly matches the share of primary commodity inputs in the U.S. CPI. For the share of commodities in the production *function,* we also use 10 percent, as in Arseneau and Leduc (2013); given that we focus on commodities rather than oil alone, both values are larger than the 5 percent commonly used in oil-only models. Finally, the size of the competitive commodity production sector relative to GDP is set at 10 percent.

Appendix B. The Dominant Commodity Exporter's Problem

For the dominant commodity exporter, the optimization problem can be written as a series of intratemporal decisions. Under the assumption that the dominant commodity exporter does not fully internalize the actions of the other exporters (i.e., taking as given the macroeconomic variables (C_t, MC_t, Y_t, Δ_t, and Ω_t) of the commodity-importing country), the problem can be written as

$$\max_{\mathfrak{M}_t} \ln\left(Q^{1/(1-\gamma)}\mathfrak{M}_t - \mathfrak{M}_t/Z_t\right)$$
$$\text{s.t. } Q_t = h(\mathfrak{M}_t).$$

The first-order condition of this problem is

$$Q_t = \left[Z_t^{-1}\frac{1}{1-\zeta^{-1}\eta_t}\right]^{\zeta},$$

where $\eta_t \equiv -\partial \ln Q_t/\partial \ln \mathfrak{M}_t = -h'\left(\mathfrak{M}_t\right)\mathfrak{M}_t)/Q_t$ is the elasticity of commodity demand (in absolute value) and $\zeta = 1-\gamma$.

The demand for the commodity by the dominant commodity producer takes the form

$$\mathfrak{M}_t = \frac{1}{Q_t}\mathfrak{D}_t - Q_t\mathfrak{E}_t,$$

where $\mathfrak{D}_t \equiv (\gamma C_t + \alpha MC_t Y_t \Delta_t)$ and $\mathfrak{E}_t \equiv \Omega_t Z_t$.

The inverse demand function is

$$Q_t = \frac{1}{2}\frac{\sqrt{\mathfrak{M}^2 + 4\mathfrak{D}_t\mathfrak{E}_t} - \mathfrak{M}}{\mathfrak{E}_t}.$$

Given this, the elasticity of demand for the commodity is

$$\eta_t = -\frac{f'\left(\mathfrak{M}_t\right)\mathfrak{M}_t}{Q_t} = \frac{\mathfrak{M}_t}{\sqrt{\mathfrak{M}_t^2 + 4\mathfrak{D}_t\mathfrak{E}_t}} = \frac{\mathfrak{M}_t}{\sqrt{\mathfrak{M}_t^2 + 4(\mathfrak{M}_t + X_t)X_t}}$$
$$= \frac{\mathfrak{M}_t}{\mathfrak{M}_t + 2X_t}.$$

Finally, the commodity price markup is

$$\Psi_t = \frac{1}{1-\eta_t} = 1 + \frac{\mathfrak{M}_t}{2X_t}.$$

Appendix C. Output, Inflation, and Interest Rate Benchmarks

C.1 Efficient Output Gap Benchmark

The benchmark output gap can be derived by substituting the equations for labor demand, labor supply, aggregate demand, and commodity demand for production into the aggregate production function. The level of output in terms of marginal costs, the dispersion of prices, productivity, and the real commodity price is

$$Y_t = \left(\frac{A_t}{\Delta_t}\right)^{1/(1-\alpha)} \left[\frac{(1-\alpha)\, MC_t\, \Delta_t}{(Q_t)^{-\gamma/(1-\gamma)} - \alpha MC_t\, \Delta_t}\right]^{1/(1+v)} \times \left(\alpha \frac{MC_t\, \Delta_t}{Q_t}\right)^{\alpha/(1-\alpha)}.$$

The log-linear approximation of the level of output, in deviations from the steady state, is[21]

$$y_t = \frac{1}{1-\alpha} a_t + \frac{\alpha}{1-\alpha}(mc_t - q_t) + \frac{1}{1+v}\Upsilon\left(mc_t + \frac{\gamma}{1-\gamma} q_t\right),$$

where $\Upsilon \equiv \left[1 - \frac{\alpha}{\mu} Q^{\gamma/(1-\gamma)}\right]^{-1} \geq 1$.

The level of efficient output, y_t^e, is defined with respect to the efficient allocation, i.e., flexible prices and no monopolistic distortions in the commodity market or in the final goods market (which implies that $Q_t^e = Z_t^{-1}$ and $\mu^e = 1$):

$$y_t^e = \frac{1}{1-\alpha} a_t + \left(\frac{\alpha}{1-\alpha} - \frac{1}{1+v}\frac{\gamma}{1-\gamma}\Upsilon^e\right) z_t,$$

[21] Note that a linear approximation of the price dispersion, Δ_t, does not appear in this equation because price dispersion is assumed to have only second-order effects on the dynamics, as shown in Benigno and Woodford (2005).

where $\Upsilon^e = \left[1 - \alpha Z^{-\gamma/(1-\gamma)}\right]^{-1}$. The relationship between Υ^e and Υ depends on the extent of monopolistic distortions. Υ^e and Υ are equal only if both markets are perfectly competitive or if the commodity is not used for production (that is, $\alpha = 0$).[22] A key difference is that commodity shocks affecting the markup do not have an effect on the level of efficient output: such output is affected only by fluctuations associated with supply shocks in the commodity market. As a consequence, a demand-driven increase in the commodity price would leave the benchmark efficient output gap unchanged. However, a negative commodity supply shock would decrease the efficient output level.

The *efficient output gap*, $\hat{y}_t^e$, which is defined as the difference between actual output and the efficient level of output, is of the form

$$\hat{y}_t^e = y_t - y_t^e.$$

C.2 Inflation Benchmarks

Expressed in log-linear terms, the equations for headline inflation and core inflation, respectively, are of the form

$$\pi_t = \pi_{Y,t} + \frac{\gamma}{1-\gamma}\Delta q_t$$
$$\pi_{Y,t} = \kappa_y \hat{y}_t^e + E_t \pi_{Y,t+1} + u_t,$$

where u_t is an endogenous cost-push shock, which is a function of both ψ_t and z_t.

In this model, the divine coincidence featured in models with exogenous commodity prices is broken. It is no longer possible to simultaneously stabilize core inflation and the welfare-relevant output gap. The tradeoff arises from the effect of commodity prices on the level of efficient output. An increase in commodity price markups generates a positive cost-push shock, which puts upward pressure on core inflation but lowers the efficient output gap.

[22] More precisely, $\Upsilon^e > (<)\,\Upsilon$, if $\Psi^{\gamma/(1-\gamma)} < (>)\,\mu$, where Ψ and μ are the markups in steady state of the commodity and final goods markets, respectively.

C.3 Interest Rate Benchmarks

The interest rate benchmarks are derived by substituting the equations for the aggregate resources constraint and the definition of the price level into the IS equation:

$$\frac{1}{R_t} = \beta E_t \left[\frac{1}{\Pi_{Y,t+1}} \left(\frac{1 - \alpha MC_t Q_t^{\gamma/(1-\gamma)} \Delta_t}{1 - \alpha MC_{t+1} Q_{t+1}^{\gamma/(1-\gamma)} \Delta_{t+1}} \right) \times \left(\frac{Y_t}{Y_{t+1}} \right) \exp\left(g_{t+1} - g_t\right) \right].$$

The *efficient interest rate* is defined in the case where commodity and final goods markets are perfectly competitive and in this model can be written as

$$r_t^e = \left(g_t - E_t g_{t+1}\right) - \left(y_t^e - E_t y_{t+1}^e\right) - \frac{\gamma}{1-\gamma}\left(\Upsilon^e - 1\right)\left(z_t - E_t z_{t+1}\right).$$

Note that the efficient interest rate does not respond to changes in the markup.

References

Akram, Q. F. 2009. “Commodity Prices, Interest Rates and the Dollar.” *Energy Economics* 31 (6): 838–51.

Anzuini, A., M. Lombardi, and P. Pagano. 2013. “The Impact of Monetary Policy Shocks on Commodity Prices.” *International Journal of Central Banking* 9 (3, September): 125–50.

Arseneau, D., and S. Leduc. 2013. “Commodity Price Movements in a General Equilibrium Model of Storage.” *IMF Economic Review* 61 (1): 199–224.

Backus, D., and M. Crucini. 2000. “Oil Prices and the Terms of Trade.” *Journal of International Economics* 50 (1): 185–213.

Benigno, P., and M. Woodford. 2005. “Optimal Stabilization Policy When Wages and Prices Are Sticky: The Case of a Distorted Steady State.” *Proceedings*, Board of Governors of the Federal Reserve System, 127–80.

Blanchard, O., and J. Galí. 2007. “Real Wage Rigidities and the New Keynesian Model.” *Journal of Money, Credit and Banking* 39 (s1): 35–65.

———. 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. Galí and M. J. Gertler, 373–421 (chapter 7). University of Chicago Press.

Bodenstein, M., C. 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 L. Kilian. 2012. “Monetary Policy Responses to Oil Price Fluctuations.” *IMF Economic Review* 60 (4): 470–504.

Carlstrom, C., and T. Fuerst. 2006. “Oil Prices, Monetary Policy, and Counterfactual Experiments.” *Journal of Money, Credit and Banking* 38 (7): 1945–58.

Caruana, J., A. Filardo, and B. Hofmann. 2014. “Post-crisis Monetary Policy: Balance of Risks.” In *Monetary Policy after the Great Recession*, ed. Javier Vallés, 217–44. FUNCAS Social and Economic Studies.

Catao, L., and R. Chang. 2015. “World Food Prices and Monetary Policy.” *Journal of Monetary Economics* 75 (October): 69–88.

Filardo, A., and M. Lombardi. 2014. “Has Asian Emerging Market Monetary Policy Been Too Procyclical when Responding to Swings in Commodity Prices?” *BIS Papers* 77 (March): 129–53.

Fueki, T., J. Nakajima, S. Oyama, and Y. Tamanyu. 2018. “Identifying Oil Price Shocks and Their Consequences: Role of Expectations in the Crude Oil Market.” BIS Working Paper No. 725.

Hamilton, J. 1994. *Time Series Analysis.* Princeton, NJ: Princeton University Press.

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

Kilian, L., and C. Baumeister. 2016. “Forty Years of Oil Price Fluctuations: Why the Price of Oil May Still Surprise Us.” *Journal of Economic Perspectives* 30 (1): 139–60.

Leduc, S., and K. Sill. 2004. “A Quantitative Analysis of Oil-Price Shocks, Systematic Monetary Policy, and Economic Downturns.” *Journal of Monetary Economics* 51 (4): 781–808.

Montoro, C. 2012. “Oil Shocks and Optimal Monetary Policy.” *Macroeconomic Dynamics* 16 (2): 240–77.

Nakov, A., and G. Nuño. 2013. "Saudi Arabia and the Oil Market." *Economic Journal* 123 (573): 1333–62.

Nakov, A., and A. Pescatori. 2010. "Monetary Policy Trade-Offs with a Dominant Oil Producer." *Journal of Money, Credit and Banking* 42 (1): 1–32.

Natal, J. 2012. "Monetary Policy Response to Oil Price Shocks." *Journal of Money, Credit and Banking* 44 (1): 53–101.

Organisation for Economic Co-operation and Development. 2012. *Competition and Commodity Price Volatility.* OECD Publications.

Rajan, R. 2015. "Competitive Monetary Easing: Is It Yesterday Once More?" *Macroeconomics and Finance in Emerging Market Economies* 8 (1–2): 5–16.

Stuermer, M. 2017. "Demand Shocks Fuel Commodity Price Booms and Busts." *Economic Letter* (Federal Reserve Bank of Dallas) 12 (14).

Sussman, N., and O. Zohar. 2016. "Has Inflation Targeting Become Less Credible? Oil Prices, Global Aggregate Demand and Inflation Expectations during the Global Financial Crisis." CEPR Working Paper No. 11535.

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

Online Appendix to "Monetary Policy, Commodity Prices, and Misdiagnosis Risk"

Andrew J. Filardo,[a] Marco J. Lombardi,[a] Carlos Montoro,[b] and Massimo Minesso Ferrari[c]
[a]Bank for International Settlements
[b]Central Reserve Bank of Peru and Professor of CENTRUM Católica, Pontificia Universitad Católica del Perú
[c]European Central Bank

D.1 Commodity Price Shock Inference Using Bayesian Learning

An extension of the model consists in allowing the policymaker to learn from real-time feedback from the economy and the effect of the policy decision. The monetary authority updates its beliefs about the nature of the shocks using Bayesian learning. By updating its initial diagnosis, the monetary authority reduces the inherent procyclicality of misdiagnoses.

The learning setup is as follows. On impact ($t = 1$), the monetary authority acts on its prior beliefs about the source of the shock. Given the economy's reaction, the monetary authority then updates its views in the form of a posterior probability about whether the economy experienced a demand shock:

$$P_1\left(D|\pi_1, y_1\right) = \frac{L\left(D|\pi_1, y_1\right) P_0\left(\mathrm{D}\right)}{L\left(D|\pi_1, y_1\right) P_0\left(\mathrm{D}\right) + L\left(S|\pi_1, y_1\right) P_0\left(\mathrm{S}\right)},$$

where $L(D|\pi_1, y_1)$ and $L(S|\pi_1, y_1)$ are, respectively, the likelihood of demand and supply shocks based on observed output and inflation and the model's implied output and inflation.[1]

[1]The likelihood function would in general incorporate all observable variables in the model. Here, we abstract from those variables that do not enter the monetary policy reaction function.

At time $t = 2$, the posterior distribution $P_1(D|\pi_1, y_1)$ is used as the (proto-)prior for the new posterior. In general, the updating follows the recursive equation:

$$P_t\left(D|\pi_t, y_t\right) = \frac{L\left(D|\pi_t, y_t\right) P_{t-1}\left(D|\pi_{t-1}, y_{t-1}\right)}{L\left(D|\pi_t, y_t\right) P_{t-1}\left(D|\pi_{t-1}, y_{t-1}\right) + L\left(S|\pi_t, y_t\right) P_{t-1}\left(S|\pi_{t-1}, y_{t-1}\right)},$$

which by recursive substitution can be written as

$$P_t\left(D|\pi_t, y_t\right) = \frac{L\left(D|\pi_t, y_t\right) P_0\left(D\right) \prod_{\tau=1}^{t-1} P_\tau\left(\pi_\tau, y_\tau|D\right)}{L\left(D|\pi_t, y_t\right) P_0\left(D\right) \prod_{\tau=1}^{t-1} P_\tau\left(\pi_\tau, y_\tau|D\right) + L\left(S|\pi_t, y_t\right) P_0\left(S\right) \prod_{\tau=1}^{t-1} P_\tau\left(\pi_\tau, y_\tau|S\right)}.$$

This produces a sequence of posterior probabilities that the monetary authority attaches to the nature of the shocks as new information becomes available and sets the policy interest accordingly:[2]

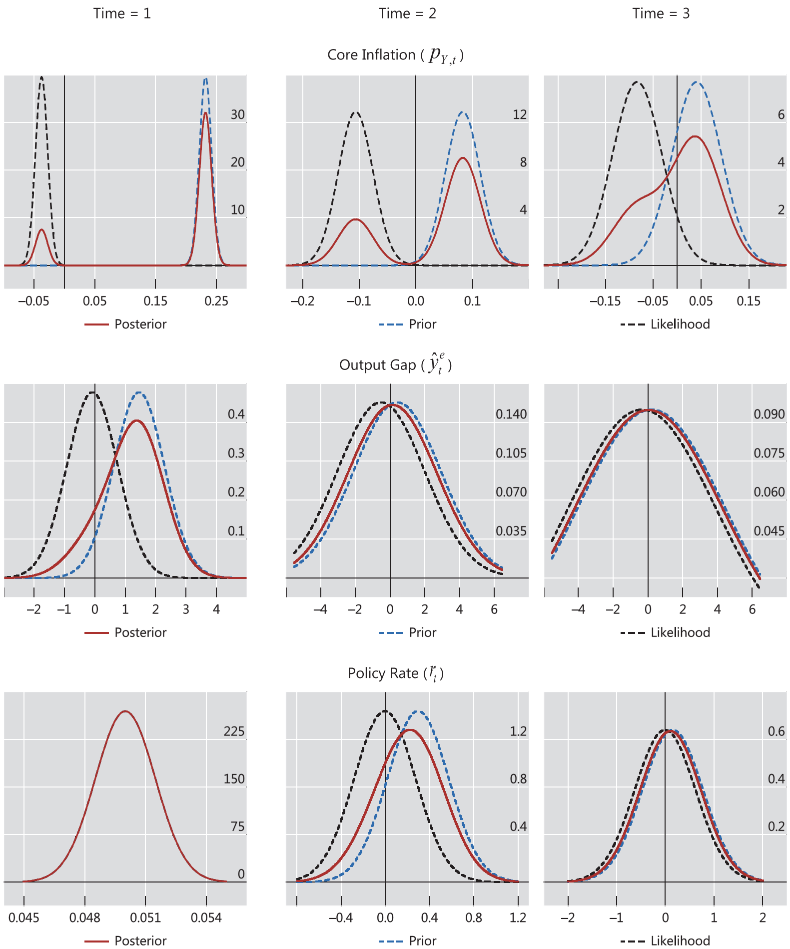
$$\begin{aligned} r_t &= E_{t|t-1}^{bayes}\left[r_t^e + \varphi_{core}\pi_{Y,t} + \varphi_y \hat{y}_t^e\right] + \varphi_{com}\Delta q_t \\ &= E_{t|t-1}\left[r_t^e + \varphi_{core}\pi_{Y,t} + \varphi_y \hat{y}_t^e\right] + \varphi_{com}\Delta q_t + e_t^{bayes}, \end{aligned}$$

with e_t^{bayes} the new misperception error. We first consider the case of a supply shock misdiagnosed as a demand shock (misdiagnosis case A in figure A.1). We start by setting the monetary authority's prior probability of a demand shock at $P_0(D) = 0.99$ and of a supply shock $P_0(S) = 0.01$.

In $t = 1$, the monetary authority observes output and inflation. Given that the authority believes the originating shock came from the demand side, the mass of the prior probability distribution is centered around the outcomes that would have occurred if the originating shock had been a demand shock (the dashed blue distribution which is calibrated to the $t = 1$ baseline responses). The likelihood reflects the information about the "true" supply shock and the endogenous monetary policy "shock" induced by the initial misdiagnosis; the likelihood is centered around the initial response

[2]Note that past "mistakes" during the learning process will influence the state of the economy. Our impulse responses propagate these mistakes over time and hence differ fundamentally from the impulse responses earlier in the paper.

Figure A.1. Misperception Case A with Bayesian Learning



Notes: Case A: A supply shock misdiagnosed to be a demand shock. Probability distributions for core inflation, the output gap, and the policy rate, as perceived by the monetary authority at three different points in time after the initial shock has occurred. "Prior" corresponds to the distribution under the demand shock initially assumed by the monetary authority (D); "Likelihood" is the probability distribution of observed variables that actually occur from a combination of the initial supply shock and the monetary policy error induced by misperception (S); and "Posterior" is the weighted average of the two, where the weights are given by the posterior probabilities assigned to D and S.

of the misperception A case. The resulting posterior distribution reflects the differences between the prior and the likelihood.

For core inflation at $t = 1$, the likelihood and the prior are very different, leading to a bimodel posterior distribution. The likelihood and prior distribution of the output gap (second row) is less precise and does not result in a bimodal posterior distribution.

At $t = 2$ and $t = 3$, the (proto-)priors and the likelihoods are closer to each other, as the monetary authority learns about the initial misdiagnosis. The posterior distributions become more unimodal. Note, however, that some of the convergence is due to strong equilibrium forces even in the case of policy "mistakes" that were seen in the earlier impulse responses.

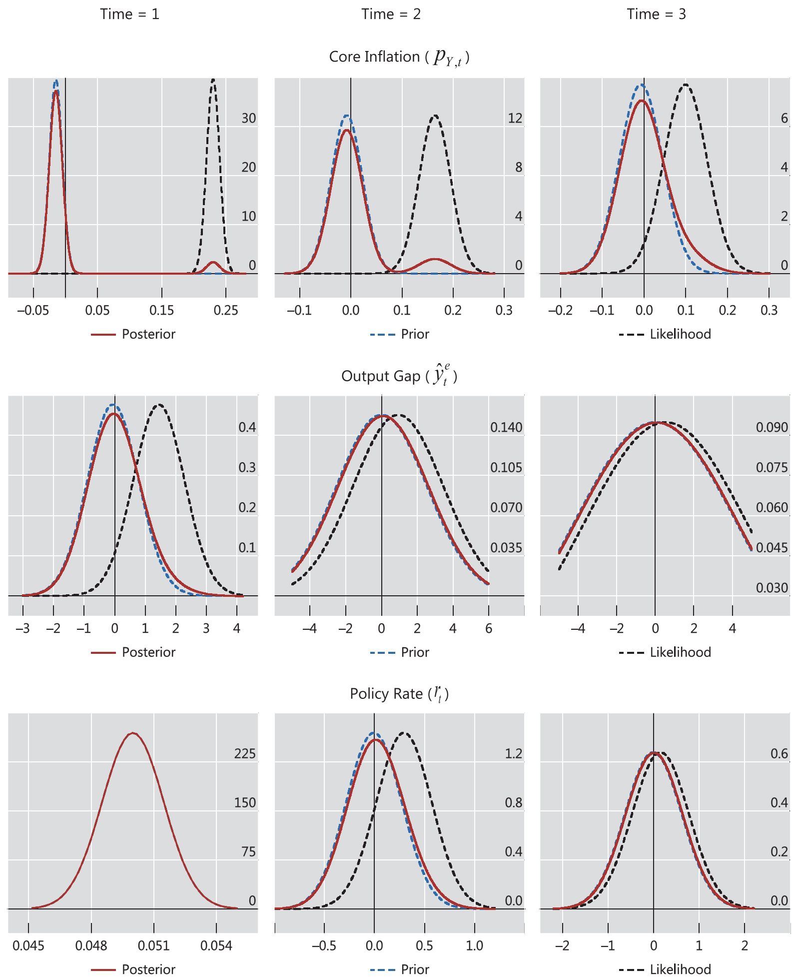
Results for misdiagnosis case B in figure A.2 indicate that the initial monetary policy response is excessively loose. As the monetary authority learns about the mistake and eventually corrects its stance, the initial procyclicality of output and inflation dies down.

This Bayesian exercise suggests that learning is a function of the shock distributions hitting the economy. Our model structure is fairly simple and generates rapid convergence via learning. However, it may be very difficult in real time for a monetary authority to converge so quickly after an initial misdiagnosis. Adding realistic shocks with fat-tailed distributions for all the variables (especially measurement error) would slow the rate of convergence and increase the persistence of the procyclical policy response. As well, parameter uncertainty (such as the extent of commodity price pass-through to inflation and the slope of the Phillips curve, which have been highlighted in recent empirical research) also complicates the inferential problem.[3]

On the other hand, if the commodity price shock is very large relative to the other shocks hitting the economy, the learning in principle should be faster, as the economy's reaction to the true shock will show through more prominently in the posterior distributions. Recent papers trying to parse the nature of oil shocks mentioned in

[3] Addressing parameter uncertainty along the lines of, e.g., Wieland (2000), Rodina (2012), and Cogley, Matthes, and Sbordone (2015) is left for future research. Additional uncertainty of this type is likely to slow the rate of learning about the nature of the commodity price shocks and make the misdiagnosis errors even more persistent.

Figure A.2. Misperception Case B with Bayesian Learning



the Introduction demonstrate just how difficult this can be empirically. And for run-of-the-mill shocks of the type calibrated in this paper, the ability of a monetary authority to learn from past mistakes during the “fog of war” suggests that lingering (endogenous) policy errors associated with misdiagnoses may be a regular feature of the monetary policy environment. Such results have far-reaching implications for all empirical macrofinancial research efforts that assume well-defined exogenous errors in the data.

References

Cogley, T., C. Matthes, and A. Sbordone. 2015. "Optimized Taylor Rules for Disinflation when Agents Are Learning." *Journal of Monetary Economics* 72 (May): 131–47.

Rodina, F. 2012. "The Role of Model Uncertainty and Learning in the US Postwar Policy Response to Oil Prices." *Journal of Economic Dynamics and Control* 36 (7): 1009–41.

Wieland, V. 2000. "Monetary Policy, Parameter Uncertainty and Optimal Learning." *Journal of Monetary Economics* 46 (1): 199–228.

The Continuing Validity of Monetary Policy Autonomy under Floating Exchange Rates*

Edward Nelson
Federal Reserve Board

Economic research in recent years has highlighted the issue of whether a floating exchange rate provides autonomy with regard to monetary policy to a central bank in an economy that is highly open. In particular, Rey (2016) has argued that inflation-targeting advanced economies lack monetary policy autonomy by pointing to results suggesting that U.S. monetary policy shocks matter for the behavior of key financial variables in these economies. In contrast, it is argued in this paper that monetary autonomy *does* prevail in inflation-targeting advanced economies, notwithstanding the reaction of these economies' asset prices to U.S. monetary policy developments. The monetary-autonomy argument—which was advanced by Milton Friedman and rests on the existence of features that are present in new open-economy models—refers to the fact that the monetary authority under a floating rate is able to have a decisive influence on nominal variables in the long run, as well as a short-run influence on real variables. The result that rest-of-world monetary policy is among the other factors affecting the short-run behavior of real variables (including real asset prices) in a small, floating-rate open economy is in keeping with the traditional and appropriate concept of monetary policy

*The author is grateful to Pierpaolo Benigno (the associate editor) and two anonymous referees for extremely useful suggestions, Sarah Baker and Laurie Khalfan for research assistance, and Ben Bernanke, Olivier Blanchard, Guy Debelle, Christopher Erceg, Joseph Gruber, Peter Ireland, Douglas Irwin, Mervyn King, David López-Salido, Ellen Meade, and George Tavlas for helpful comments on earlier drafts. None of these individuals is responsible for any errors. The views expressed here are the author's alone and should not be interpreted as those of the Federal Reserve or the Board of Governors. E-mail: Edward.Nelson@frb.gov.

autonomy under floating exchange rates. It follows that such influences of rest-of-world monetary policy on the home economy are consistent with the celebrated open-economy trilemma.

JEL Codes: E51, E52, F41.

1. Introduction

> "Keynes could state the issue as a dilemma .... [Subsequent history] has forced those of us who have written on this subject more recently to expand Keynes's dilemma to a trilemma. A country is compelled to choose two of the following three desirable objectives: stable prices (or, more generally, an independent monetary policy), a stable exchange rate (or, more generally, a predetermined path of exchange rates), freedom from exchange controls."
>
> Friedman (1983, p. 37)

Economic research in recent years has given considerable prominence to the issue of whether a floating exchange rate provides, to a central bank in an economy that is highly open, autonomy with regard to monetary policy.[1] If this autonomy existed, it would mean that the central bank—like its counterpart in a closed-economy or large-economy setting—could use monetary policy to choose a particular long-run inflation rate (perhaps in conjunction with the pursuit of other macroeconomic goals, such as keeping output close to potential). As the foregoing quotation suggests, a standard result in monetary economics is that autonomy of this kind does arise from a floating-exchange-rate arrangement and that the autonomy prevails even in conditions of complete international mobility of capital.[2]

[1] This prominence is evidenced by much of the material in Bordo and Taylor (2017).

[2] In the present discussion, complete or perfect capital mobility refers to a situation in which official (that is, governmental) controls on private-sector access to foreign exchange are not present. When reference is made here to complete capital mobility, it is not being taken for granted that capital markets in the home economy or foreign economy are free of imperfections or frictions or that agents in the small economy can obtain funds from the foreign economy on exactly the same terms as those available to a comparable agent in the foreign economy.

This result is often contrasted with the situation facing a small open economy when its exchange rate is fixed. In the case of a completely fixed exchange rate and globally mobile capital, monetary policy in the small open economy is directed toward stability of the external value of the currency and cannot be used in pursuit of objectives distinct from that goal.[3]

The conclusion that monetary policy autonomy can be secured by a floating-rate arrangement has been challenged by Rey (2013). On the basis of empirical evidence, Rey argues that, under unimpeded international capital mobility, *neither* fixed exchange rates *nor* floating exchange rates are associated with monetary autonomy for a central bank. She contends that what many economists—including Milton Friedman in the above quotation—have called the "trilemma" is invalid.[4] Even with a flexible exchange rate, she suggests, monetary autonomy would be obtainable only if the open economy's authorities restricted capital account transactions.

Although Rey (2013) focused on emerging economies, her basic argument concerning floating rates would, if valid, apply to emerging economies and advanced economies alike. And indeed Rey (2016) has extended the argument to small- and medium-sized

---

Indeed, in many new open-economy models, including some of those discussed below, it is assumed that citizens of a small country who borrow from abroad encounter some form of fee or spread that makes their overall international borrowing cost different from the rest-of-world interest rate. Such model features may be appealing for technical reasons (see Schmitt-Grohé and Uribe 2003) and also provide a more realistic account of the borrowing situation facing many small economies, including those that impose no exchange controls. But the inclusion of these model features does not in itself imply a violation of the assumption of complete mobility of international capital.

[3]Under a fixed exchange rate, monetary developments will still be decisive for the long-run inflation rate, as stressed in McCallum (1996, p. 143). However, with monetary policy choices dictated by the commitment to fix the exchange rate, monetary policy at home will be closely connected to rest-of-world monetary policy, and the economy's long-run inflation rate will consequently tend to be driven by that of the rest of the world.

[4]Obstfeld (1998, p. 14) credited the use in the research literature of the "trilemma" terminology (to describe the conventional wisdom that monetary policy autonomy, fixed exchange rates, and complete capital mobility are not jointly obtainable, though any two are) to Obstfeld and Taylor (1998). This terminology was, however, used in the same context much earlier by Friedman in the press article quoted at the start of this paper, and this appears to have been the first published instance of such usage.

advanced economies.[5] Accordingly, the discussion that follows focuses on advanced open economies—in particular, inflation-targeting countries that float their exchange rate.

The monetary-autonomy debate is of great relevance to such economies because both these economies' policymakers and outside observers have routinely accepted the standard argument in favor of autonomy. The United Kingdom, for example, has had a floating exchange rate since September 1992. Shortly after this float began, the Chancellor of the Exchequer, Norman Lamont, proclaimed: "We are floating and we will set monetary policy in this country... It will be a British economic policy and a British monetary policy."[6] The specific monetary policy chosen by Lamont—and continued over the past two decades of Bank of England operational independence—consisted of a strategy of inflation targeting. A second example is provided by Australia, which has had a floating exchange rate since 1983 and an inflation-targeting monetary policy strategy since the 1990s. Beaumont and Cui (2007, p. 1) stated that this experience had been associated with Australia "gaining the expected macroeconomic benefits from exchange rate flexibility," including monetary policy autonomy. The literature critical of the connection between floating and monetary policy autonomy is therefore a major challenge to the conventional wisdom concerning inflation-targeting economies.

In contrast to the conclusion of that literature, the position advanced here will be that a floating exchange rate *does* secure monetary policy autonomy. Consequently, provided that the exchange rate floats, official controls on capital mobility are not necessary to secure such autonomy.[7] Monetary policy autonomy under flexible exchange rates, it will be argued, is not merely an analytical result.

[5]The term "advanced economies" below will generally be used to refer to these small- and medium-sized floating-rate advanced economies, whose central banks target inflation. Likewise, in the discussion in this paper, the phrase "home economy" or "open economy" will often stand in for "small open economy." Also, the terms "capital controls" and "foreign exchange controls" will be used interchangeably, while "reserves" will refer to commercial banks' reserve balances (a central bank liability) and not to foreign exchange reserves (a central bank asset).

[6]Quoted in *The Economist* (1992).

[7]This does not, of course, preclude the validity of other possible justifications for imposing such controls. See, for example, Pasricha (2017) for an analysis of different motivations for capital controls.

Rather, it has practical validity and aids understanding of monetary and economic behavior in inflation-targeting economies. Results advanced by Rey and others in support of the contrary position do not, in fact, provide persuasive evidence of lack of monetary autonomy. The reason those results are not persuasive evidence is that the existence of monetary autonomy does not preclude, and indeed is highly consistent with, international financial integration.

The Basis of the Counterargument. A quotation from the abstract of Rey (2016) is helpful in bringing out the counterargument advanced here. Rey states: "The paper presents evidence that U.S. monetary policy shocks are transmitted internationally and affect financial conditions even in inflation-targeting economies with large financial markets. Hence flexible exchange rates are not enough to guarantee monetary autonomy in a world of large capital flows." One could not ask for a clearer articulation of the no-autonomy argument than the two sentences just quoted. But referring to those same sentences also provides a convenient means of expressing the crux of the counterargument, which is simply: The conclusion given in the second sentence does not follow from the first. That is, the word "Hence" connecting the sentences is unwarranted and the second sentence is a non sequitur.[8]

In fact, the transmission of U.S. financial developments (including those arising from U.S. monetary policy shocks or other U.S. monetary policy actions) to financial conditions in other, smaller, advanced economies is fully consistent with the possession of monetary autonomy by the central banks of those economies. The key point, as already suggested, is that monetary policy autonomy under a float can, and likely does, coexist with financial interdependence.

Objectives of the Present Analysis. Elements of the preceding point can be gleaned from some of the critical discussions of Rey (2013). But, to date, the communication of that point has been weakened by being left implicit, by taking the form of a side remark, or by appearing in the context of discussions that make unwarranted

[8]The approach taken here is therefore consistent with, but somewhat different from, that in Taylor's (2016) defense of monetary autonomy. Taylor focuses on whether the empirical findings offered in recent years against autonomy are artefacts of the estimation sample periods, rather than on whether—if subsequently found to be durable—those results actually point to the absence of autonomy.

concessions to the no-autonomy position.[9] A first objective of this paper, therefore, is to make the point central and explicit, doing so with specific reference to advanced inflation-targeting economies.

A second, related, objective of this paper is to bring out the shortcomings of the recent case against autonomy under floating exchange rates by comparing this case directly with the *standard* case for floating exchange rates. It will become clear that the results claimed to be at variance with the existence of monetary autonomy do not actually conflict with the basic notion of monetary autonomy advanced by advocates of floating rates.

In particular, the discussion in this paper will consider the monetary-autonomy argument used in Milton Friedman's (1953a) case for exchange rate flexibility. That paper's argument is a useful benchmark because Friedman (1953a) is widely accepted as a central reference on the issues discussed here.[10] Bringing his argument explicitly into the current debate highlights the sharp difference between (i) what believers in monetary autonomy under a floating rate thought was implied by autonomy and (ii) what modern researchers have regarded as evidence that autonomy does not prevail under floating rates.

It has been rare for the recent literature on monetary autonomy to cite Friedman's views on the matter. For example, Rey (2013, 2016) does not cite Friedman's work. And when it has characterized Friedman's autonomy argument, the recent literature has sometimes attributed to him the opposite of his actual position.[11]

Rey (2016) does cite the classic work of Mundell (1963) that provided formal analysis of the incompatibility, for a country allowing complete capital mobility, of fixed exchange rates and an autonomous monetary policy. However, when considering the debate on monetary autonomy, Friedman's contributions provide a

---

[9] A couple of prominent examples of these studies are considered in section 2.4 below.

[10] For example, Krugman (1993, p. 519) judged Friedman (1953a) to be a "seminal paper," while McCallum (1996, p. 213) observed that Friedman (1953a) was "the most famous and influential single piece of writing on exchange rate arrangements." See also Dellas and Tavlas (2018) and Irwin (2017) for analyses of the development of Friedman's article and the historical context in which it appeared.

[11] See the discussion in section 2.3 below of Obstfeld and Taylor's (2017) portrayal of Friedman's argument.

starker and more apposite counterpoint to Rey's position than does Mundell's work. There are three reasons for this. First, although Rey calls the trilemma the "Mundellian trilemma," the "trilemma" terminology was used in print by Friedman more than 30 years ago, long before it became prevalent; and as discussed below (in section 2.2), he articulated the trilemma concept when laying out the case for floating rates in 1953. Second, as detailed later, Rey materially misstates the properties of the model Mundell developed, as she attributes it to a feature (the central bank's pursuit of stabilization policy using an interest rate instrument) that is infeasible in that model (under both fixed and floating rates) but that is possible, under floating rates, in Friedman's framework.[12] Third, Mundell himself became a strong critic of floating exchange rates (see, for example, Mundell 1968) and so, unlike Friedman, he is not particularly representative of the view that floating rates (alongside capital mobility) are attractive from the point of view of stabilization policy.

In the course of the analysis that follows, it will be shown that the autonomy obtainable under flexible exchange rates in sticky-price models lines up with the autonomy Friedman described in his writings. They have a common implication: asset prices can move together across countries whose exchange rates float, yet this does not imply that an open economy lacks monetary autonomy. A corollary is that empirical evidence that asset prices in inflation-targeting advanced economies respond to U.S. monetary policy actions does not in itself constitute valid evidence against autonomy.

The focus here is therefore on the generic problems with the type of evidence offered of late against autonomy under floating rates. This paper will not review that evidence in detail. Indeed, the empirical findings per se will not be disputed at all.[13] What will be challenged is the inference that such findings are evidence against the notion that monetary policy autonomy prevails under floating exchange rates. In contrast to Rey's (2016, p. 27) suggestion that

[12]The Mundell-Fleming model does feature monetary autonomy under floating rates. But the autonomy is of a kind that is of questionable practical relevance, because it is one under which the central bank is incapable under floating rates (alongside capital mobility) of managing interest rates.

[13]Nor will there be systematic discussion of the empirical evidence offered *in favor of* autonomy in Klein and Shambaugh (2015) and Obstfeld, Ostry, and Qureshi (2017), for example.

"many more VARs need to be run" before the hypothesis of lack of monetary autonomy can be accepted, the perspective of the present paper is that the hypothesis would not be valid even if Rey's finding of effects of rest-of-world monetary policy shocks on domestic variables is fully granted.

This paper proceeds, in section 2, by considering the standard argument for monetary policy autonomy under floating exchange rates, with a focus on Friedman's (1953a) exposition of the argument. Relevant characteristics of new open-economy models under fixed and floating rates are then discussed in section 3. Section 4 reconsiders, in light of the analysis of the preceding sections, the broad evidence presented against autonomy. Section 5 concludes.

2. The Standard Argument for Monetary Policy Autonomy under Floating Exchange Rates

Before proceeding further, a specific characterization of monetary autonomy, and the argument that supports it, is needed.

What precisely is the monetary policy autonomy that floating rates should provide to a small open economy under full capital mobility? The answer—according to the standard case for a floating exchange rate—can be expressed in terms of *aggregate macroeconomic outcomes* as follows: In both the short run and the long run, the central bank can have a decisive influence on the monetary value of economic activity, that is, the aggregate nominal flow of spending on goods and services in the economy.[14] This influence

[14]Expressing the central bank's influence in terms of ability to affect nominal spending may appear counterintuitive, as nominal income is typically not a variable that appears in the structural equations of closed-economy and open-economy models, whereas its individual components—prices and real output—do so. However, it is highly convenient to refer initially to nominal spending because this is a variable that central banks can, under monetary autonomy, affect *at all time horizons* and *across a variety of models*. In particular, under monetary autonomy, an open economy's central bank is able to set the course for nominal spending, in both the short run and long run, even if it cannot affect real output (as is the case at all horizons in flexible-price models, and in the long run of sticky-price new open-economy models) or cannot set the inflation rate (either because domestic prices and costs are rigidly fixed at all horizons—as they were in some very early open-economy models—or because the home country's inflation rate is heavily influenced by nonmonetary factors even in the long run—a

can be further categorized in terms of the short-run situation and long-run conditions. In the short run, monetary autonomy, in combination with price stickiness, gives the monetary authority considerable influence over both real and nominal variables. In the long run, monetary autonomy gives the monetary authority commanding influence over nominal variables, even though its influence on real variables is gone (because price stickiness wears off over time, so real variables return to their natural values). The influence over the long-run nominal situation, together with the acceptance by policymakers that they have this influence, has prompted many open economies' monetary authorities to select and pursue a target for their country's long-run inflation rate.

This influence on macroeconomic *outcomes* in turn stems from the fact that, as detailed below, the central bank under a float is able to set a course for its policy *instruments* that does not depend mechanically or solely on the behavior of the exchange rate or on developments abroad.

With that background established, the remainder of this section shows how this concept of autonomy relates to Friedman's (1953a) case for floating exchange rates (section 2.1) and his expression of the trilemma (section 2.2). Then section 2.3 reconsiders the critiques of the autonomy argument in recent literature, while section 2.4 describes the relationship between the defense of the trilemma argument made here with two other recent defenses.

2.1 Friedman's (1953a) Emphasis on Monetary Policy Autonomy

The autonomy argument described above was an important component of the Friedman (1953a) case for floating exchange rates alongside full capital mobility.[15] With regard to *macroeconomic outcomes*,

position obviously contrary to the modern consensus, but one apparently held by some advocates of floating rates, such as Meade 1961).

[15]The present discussion will not consider Friedman's case for floating exchange rates in toto but, rather, will focus on the monetary-autonomy aspect of that case. A number of aspects of Friedman's case for floating rates (such as his predictions concerning whether stable monetary policies could be counted on to generate fairly stable exchange rates and his contentions about the relationship between speculation and exchange rate stability) have arguably not been borne out by

Friedman stressed that monetary policy autonomy implied that the domestic monetary authority had a decisive influence on the behavior of nominal economic aggregates. Thus, floating exchange rates allowed a country to achieve "monetary stability" and to set the criterion for such stability (p. 200). In particular, Friedman stressed that each individual country separately under floating exchange rates can achieve "avoidance of either inflation or deflation," and likewise "any one country" can follow a policy of inflation or deflation without that policy choice being imposed on, or inherited by, other countries (Friedman 1953a, pp. 198, 199).

Furthermore, under fixed exchange rates, balance of payments deficits could occur and act as a negative influence on nominal aggregate demand. With prices sticky in the short run, this would create situations of deficient real aggregate demand and above-normal unemployment. Flexible exchange rates, in contrast, implied a zero overall balance of payments. This zero balance in turn made it possible for the domestic monetary authority to make policy decisions that implied a particular path for nominal variables, such as nominal aggregate spending on goods and services. In the short run, this power also gave the monetary authority the ability to influence real aggregate demand, an ability that it might use to pursue full-employment goals (Friedman 1953a, pp. 165–167, 171).

These influences on macroeconomic outcomes—including the long-run behavior of inflation and the short-run behavior of output and other real variables—were in turn traceable to the fact that monetary autonomy gave the central bank prerogatives over the settings of the country's monetary *instruments*. Friedman primarily articulated this influence on instruments in terms of monetary quantities. In this vein, Friedman (1953a, pp. 181, 200, 201) referred to a monetary authority possessing autonomy as able to "create ... money,"

events; yet the monetary-autonomy argument may remain valid in the face of such invalid aspects of the Friedman case. Along the same lines, one can accept Obstfeld and Taylor's (2017) contention that sizable and fluctuating capital flows have been an enduring part of the environment that floating-rate countries face—and to a far greater degree than Friedman envisioned in 1953; yet one could also view flexible exchange rates as a means of securing monetary autonomy in the presence of such capital account fluctuations. Indeed, in his later expositions, in which he acknowledged the continuing volatility of capital flows, Friedman reaffirmed monetary autonomy as an advantage of floating rates (see, for example, Friedman and Friedman 1984).

achieve "currency issue," and undertake "[m]onetary expansion." Noting that countries that have domestic macroeconomic objectives, such as full employment and price stability, would seek "control over domestic monetary policy" (p. 180), he observed that such control could be achieved under arrangements in which the rest-of-world developments were not an automatic influence on the domestic monetary base and money stock (p. 199). But, under free trade and full capital mobility, such arrangements were obtainable only under floating exchange rates. Consequently, Friedman regarded a floating rate as conferring to a central bank control over the monetary base and, with that control, also a decisive influence over the money stock.

It is, however, readily possible to express this autonomy in terms of interest rates instead of monetary quantities. As the issuer of the liabilities through which interbank transactions occur, and as an entity that has a commanding influence over the demand/supply intersection in the reserves market, the central bank strongly influences the interest rate at which interbank lending takes place. Thanks to links between different financial markets, such an influence on short-term interbank rates translates into an influence on domestic short-term rates more broadly. If, however—as discussed further below—the central bank is enforcing a fixed exchange rate under conditions of full capital mobility, it in effect cedes this influence on short-term interest rates, because it is obliged to make rates at home move in lockstep with comparable rates abroad.

It is thus clear that, even in the modern day—in which a central bank's influence on the supply of monetary base tends to be deemphasized in monetary analysis—Friedman's focus on the consequences for monetary control of different exchange rate arrangements remains vital. That this is so is brought out by the fact that, early in the era of U.K. inflation targeting, King (1994, p. 268) noted that a central bank is well positioned to manage short-term market interest rates whether it operates primarily by changing the volume of commercial banks' reserve balances or by actions that shift the demand curve for bank reserves.[16] His argument *took for granted*

[16]This was the case even when Friedman wrote in the early 1950s, when one tool available to central banks was the reserve-requirement ratio, variations in which could shift commercial banks' demand curve for reserves.

the existence of a floating-rate regime. For under a fixed exchange rate and full capital mobility, the interest rate in the market for overnight debt must take the value implied by the exchange rate target, and the central bank must acquiesce in conducting operations that deliver that interest rate. This is true irrespective of the operating procedure of the central bank. Therefore, even when the central bank manages short-term market interest rates primarily by altering commercial banks' demand for reserves (via an interest-on-reserves policy, for example), floating exchange rates are required if the central bank's interest rate management is to be used for domestic stabilization purposes. In this connection, Woodford (2010, pp. 43–46) analyzes the case in which the central bank manages domestic short-term market interest rates by varying the interest rate it pays on bank reserves. His analysis takes place under the assumption of a floating exchange rate.[17] That assumption implies that the central bank decisionmaking power over interest rates, derived by Woodford (2001) in a closed-economy setting, carries over to an open-economy environment.

And Friedman himself on occasion pointed to the fact that a float could be viewed as allowing the central bank to choose interest rates for the home economy. For example, in his 1953 article, he noted that key domestic interest rates were "susceptible to direct influence by the monetary authorities" (Friedman 1953a, p. 166). He thus acknowledged that monetary autonomy entailed the opportunity to influence such rates—opening the possibility of management of interest rates to achieve domestic goals.

[17]That, under fixed exchange rates, a central bank must allow the domestic short-term rate to move continuously in step with the rate abroad is an implication of the uncovered interest parity (UIP) condition. Under that condition, there is a one-for-one relationship between domestic and foreign short-term interest rates when the expected next-period change in the nominal exchange rate is zero. In terms of domestic securities markets, the forces pushing the home rate in the same direction as the foreign rate remain present when the former rate is managed using an interest-on-reserves arrangement. This can be seen by examining the two methods of managing market interest rates analyzed in Keister (2012). Both methods he considers involve paying interest on reserves, but both also involve the central bank "setting the supply of reserve balances"—something it is not at liberty to do under a fixed exchange rate. Instead, when the exchange rate is fixed, international payments flows will steer reserve balances in a direction that makes the home short-term interest rate move in tandem with the rest-of-world rate.

2.2 Friedman's Analysis of the Trilemma

Friedman's (1953a) case for monetary autonomy under floating exchange rates also amounted to an argument for the existence of the trilemma. Friedman's first public use of the term "trilemma" was evidently in the Friedman (1983) quotation that opened this paper.[18] However, he used the term "trilemma" in an early draft version of Friedman (1953a) (see Irwin 2012, p. 35), and the trilemma concept permeates the published discussion in Friedman (1953a). The 1953 article noted that, in the postwar world, foreign exchange controls were a means by which countries, like the United Kingdom, that were committed to implementing demand-management policies had tried to reconcile fixed exchange rates with monetary autonomy (pp. 169, 179). Friedman opposed this third alternative of fixed exchange rates plus exchange controls because of the inefficiency that the controls engendered and because the private sector would likely find ways to evade any fixed set of regulations regarding international transactions (pp. 163, 169). Such parts of Friedman's (1953a) analysis led Dornbusch (1986, p. 31) to judge that the article was a benchmark reference not only for the case against fixed exchange rates but also for the case against exchange controls.

2.3 What Friedman Did Not Say

In the context of a discussion of monetary autonomy, it is also important to be clear on what Friedman did *not* say. Such clarity is needed because recent discussions have associated with the monetary-autonomy argument positions that do not, in fact, follow from that argument. The account below of what Friedman did not say proceeds by analyzing what was attributed to the monetary-autonomy argument by Rey (2016) and by Obstfeld and Taylor (2017).

- *Rey (2016):* Rey (2016, p. 7) implies that monetary autonomy requires that a float "insulate[s] the domestic economy ... from global factors" However, Friedman's (1953a) concept

[18] As already implied, the 1983 Friedman article also appears to have been the origin in print of the use of the term "trilemma" to describe the opportunities available to an open economy's central bank.

of autonomy did not entail such comprehensive insulation. Friedman stated explicitly that, under floating, international shocks would still tend to affect real variables in the domestic economy. He contended that, while developments abroad would continue to matter for the home economy under floating exchange rates, a float gave the central bank the option of preventing those developments from operating via "monetary channels," and that "monetary stability" was still possible for a country whose exchange rate floated, irrespective of the monetary policy pursued abroad (Friedman 1953a, p. 200).[19] What monetary autonomy entailed was the absence of the "domination of internal monetary policy by external forces" (in the words of Friedman, 1953b, p. 203)—not the absence of any dependence of the economy on external forces.

In Federal Reserve Bank of Kansas City (2013, p. 353) Rey said: "We think of the trilemma as[:] if you have a fully free floating exchange rate, you are completely insulated from the external. That has been my interpretation of a trilemma." As already indicated, however, the standard case for floating exchange rates, along with the implied concept of the trilemma, makes no such wide-ranging claim about the implications of floating exchange rates.

- *Obstfeld and Taylor (2017)*: Obstfeld and Taylor (2017, p. 12), in discussing professional views on exchange rates during the Bretton Woods fixed-exchange-rate era, have stated:

> More academic economists began to echo the early calls by Friedman (1953[a]) ... for floating exchange rates, arguing that market-determined rates would tend to eliminate external payments imbalances while insulating countries from foreign inflationary shocks. Their basic argument was that routine exchange-rate flexibility allows

[19]In Friedman and Roosa (1967, p. 104), Friedman elaborated that a floating rate did not insulate an economy from those "external events that do require changes in the pattern of production and consumption." He observed that, in these cases, a floating rate allowed the domestic economy to undergo "adjustment to the change in real factors" abroad, albeit without (permanent) aggregate price-level movements being part of the adjustment process. See also Friedman (1953a, p. 182).

> all countries to move to a preferred resolution of the trilemma—as compared with the situation of much more constrained policymaking that they then faced. As Johnson (1969, p. 18) put it: "Flexible rates would allow each country to pursue the mixture of unemployment and price trend objectives it prefers, consistent with international equilibrium, equilibrium being secured by appreciation of the currencies of 'price stability' countries relative to the currencies of 'full employment' countries."

The preceding passage surely implies that the Johnson (1969) quotation accurately summarizes the Friedman concept of monetary autonomy. Yet the Johnson quotation is definitely predicated on the existence of a permanently downward-sloping Phillips curve: a permanent state of affairs in which underemployment buys price stability, and in which full employment can be obtained provided that inflation is permitted. But that is not the *Friedman* position.[20] Friedman was emphatically not a subscriber to the belief that the Phillips curve was permanently downward sloping, yet it is certainly implied by Obstfeld and Taylor (2017) that he was. It is clear that Friedman did not see floating rates as leaving policymakers free to select a long-run combination of inflation and unemployment from a Phillips-curve menu. Rather, the long-run freedom conferred by floating rates in Friedman's vision pertained only to the choice of the inflation rate.[21]

[20] The Johnson (1969) paper was apparently intended in large part to "improve" upon Friedman (1953a) by adding to the case for floating rates an argument drawing upon then-prevalent beliefs that the Phillips curve was permanently nonvertical. In such circumstances, a monetary authority could accept some departure from price stability to obtain full employment, run still-higher inflation to keep the economy in an overemployment situation, or choose price stability at the cost of permanent resource slack. However, such beliefs went out of favor within a few years of the appearance of Johnson's paper, in part because of Friedman's own critique of the Phillips curve. Friedman's belief in the compatibility of full employment and price stability dated back to before the appearance of his 1953 essay. See Nelson (2009, pp. 44, 70) for a detailed discussion of these matters.

[21] Friedman *did* see floating rates as making policymakers better placed to stabilize employment in the short run (see Friedman 1953a, p. 158, and the passages already noted). Such a stabilization role for monetary authorities in no way implies an ability on their part to choose the long-run level of employment.

In addition, the phrase "eliminate external payments imbalances" in the above Obstfeld-Taylor passage captures Friedman's position only if the "payments imbalances" in question are those for the aggregate balance of payments, not its current account and capital account categories considered individually. Friedman argued that a floating rate would make the overall balance of payments zero. Nonzero current account imbalances (matched by nonzero capital account imbalances) under floating rates are wholly compatible with Friedman's argument.

Obstfeld and Taylor (2017, p. 15) further state: "Early advocates of floating exchange rates like Friedman . . . clearly oversold the extent to which they could facilitate trade while still insulating a domestic economy from international shocks." Not only did Friedman not oversell the position attributed to him in this quotation, he did not even subscribe to that position. As indicated above in the discussion of Rey (2016), the Friedman position was *not* one in which a floating exchange rate makes the domestic economy wholly insulated from international shocks.

Friedman confined the macroeconomic variables that could, in the long run, be insulated from foreign influences under floating rates to nominal economic series—the control of which flowed from the central bank's ability to set the course for monetary instruments. The central bank's power to adjust policy instruments also raised the possibility that the central bank could exercise a temporary influence on real series in the home economy. In particular, as indicated above, Friedman saw short-run stickiness of prices as bequeathing to monetary policy the ability to influence the course of real variables in the short run. Monetary policy then might, but need not, attempt to offset the effects of international real shocks on domestic real variables in the short run. But, due to the temporary nature of price stickiness, even a monetary policy that attempted to prevent the short-run effects of real foreign shocks on domestic economic activity would not be capable of stopping real factors—including persistent foreign real shocks—from being decisive for the long-run behavior of real variables.

Part of this foreign influence on the home economy under floating exchange rates is on (and via) asset prices in the home economy. This asset price response is one consequence of international financial integration. International financial integration is a phenomenon that needs to be sharply distinguished from monetary policy autonomy—as will be stressed in section 4 below, when Rey's (2016) empirical results are reconsidered.

Obstfeld and Taylor (2017) conclude in favor of the trilemma, but, as the preceding quotations indicate, they give the impression that Friedman's reasoning in support of the trilemma was mistaken and that the concept can only be salvaged using different, later arguments. Indeed, they specifically state that Friedman "erred" (p. 15) and offer as the correct position the point that "when faced with external shocks, countries with floating exchange rates still have a shock absorber that countries that peg exchange rates lack and thus can achieve preferred policy outcomes even if they cannot achieve full insulation of their economies (Obstfeld 2015)."[22] Their attribution of this point solely to Obstfeld (2015), a paper that does not cite Friedman, alongside the authors' negative judgment on Friedman's argument, has a clear implication: that the limited-insulation, "shock absorber" notion is one that supersedes and replaces the argument for monetary policy autonomy under floating rates given in Friedman (1953a). However, as already indicated, Friedman did not make the claims attributed to him by Obstfeld and Taylor (2017), and his vision of the exchange rate's role actually mirrors the "shock absorber" function that they attribute to Obstfeld (2015). Indeed, Friedman himself on occasion used the "shock absorber" phrase for his conception of a floating exchange rate (see, for example, Friedman 1975).

In sum, in making the case for the ongoing importance of the trilemma and the continuing validity of monetary policy autonomy under floating exchange rates, one can rely on the standard argument for monetary policy autonomy, as outlined in particular in Friedman (1953a).

2.4 Comparison with Recent Endorsements of the Trilemma

It is worthwhile to put the reaffirmation of the validity of monetary autonomy made in this paper in the context of two recent endorsements of the trilemma. This comparison brings out areas of agreement but also highlights the respects in which the case made here

[22] From Obstfeld and Taylor (2017, p. 17).

goes further than existing endorsements. The discussions considered are those of Bernanke (2017) and Debelle (2017).

- *Bernanke (2017):* Bernanke endorses the trilemma and the implication that monetary policy autonomy is obtainable under exchange rate flexibility, even when complete capital mobility connects countries' financial systems (see Bernanke 2017, p. 7). But much of Bernanke's subsequent discussion is concerned with the circumstances under which a central bank decides to assign to monetary policy tasks other than stabilization of domestic economy-wide variables—not with whether stabilization is an available option.

  Bernanke (2017, p. 15) indicates that, in the model he considers, the use of monetary policy in a way that makes output equal potential output is an option. He therefore challenges the claim of no autonomy and reaches a result highly consistent with the analysis given above. However, this result is presented in the course of an analysis in which the exchange rate is "the policy control variable" (p. 15), when in fact a defining characteristic of targeting domestic variables under floating rates is that the exchange rate is not controlled.

  Bernanke also likens monetary policy autonomy to results obtained under flexible exchange rates "in the standard Mundell-Fleming analysis" (p. 15). However, in the standard Mundell-Fleming analysis the domestic interest rate (but not the money stock) is pinned down by world conditions even under floating rates, provided that capital is internationally mobile. On that dimension, the Mundell-Fleming analysis gives results unlike those that one would normally associate in practice with monetary policy autonomy, as under autonomy one would expect key domestic interest rates to be affected by actions of the home country's central bank. Consequently, a parallel with Friedman (1953a), as above, or with new open-economy models (as in section 3 below) would seem more germane, as these analyses, unlike Mundell's, consider the case in which a float confers on the central bank the ability to choose values for domestic interest rates.
- *Debelle (2017):* Debelle notes that the Reserve Bank of Australia (RBA) controls a short-term interest rate and argues

that management of this rate can deliver the aggregate-demand and inflation outcomes sought by the RBA in its inflation-targeting strategy. His argument amounts to a denial of the no-autonomy position. But this fact is largely left implicit. Indeed, Debelle (2017) states:

> Rey has recently described this state of affairs as a monetary dilemma. . . . That is, we can only set the monetary policy we want if we impose controls on the flow of capital in and out of the country. I don't think the situation is quite as stark as that. There is still a substantial degree of flexibility to set domestic monetary policy appropriately for domestic conditions. But I would certainly agree that the monetary policy decisions of other central banks are a significant factor to be taken into account in our monetary policy deliberations. Another way of stating this is that we don't have the independence to set the neutral rate, which is significantly influenced by global forces, but we do have the independence as to where we set our policy rate relative to the neutral rate.

Although the conditions stated that begin with the sentence starting "But. . . " are presented as though they partially reconcile the Debelle and Rey positions on the situation facing a central bank under floating rates, they do not in fact imply any true concession to the Rey position of lack of monetary autonomy. International factors, including foreign monetary policy, can matter for the evolution of the domestic aggregate variable(s) targeted by a central bank that has monetary autonomy. And taking the neutral interest rate as externally given is a situation that a central bank even in a large or closed economy typically faces. Therefore, for a small open economy, the fact that the central bank does not set the neutral rate does not imply an absence of monetary policy autonomy.

3. Floating Exchange Rates and Monetary Autonomy in Different Model Environments

The conditions associated with monetary policy autonomy that are predicted by the standard argument of Friedman (1953a) are present in formal dynamic models of the open economy. This is brought out later, when the properties of sticky-price new open-economy models under a float are discussed. As a preliminary matter, however, it is useful to consider the operation of floating exchange rates in flexible-price new open-economy models.

3.1 Floating Exchange Rates and Monetary Autonomy in Flexible-Price Models

Instantaneous full price flexibility is not the environment to which Friedman's (1953a) description of monetary autonomy applied. But the flexible-price case gives insights into the coexistence of international financial integration and national monetary policy autonomy under floating exchange rates.

Flexible prices in the home economy mean that the home monetary authority cannot affect the short-term real interest rate or any other real variable. In addition, home-economy short-term real interest rates will likely move closely with the rest-of-world rate. As is true of both sticky-price and flexible-price models, the uncovered interest parity (UIP) condition, once expressed in real terms, indicates that the spread between the real short-term interest rate and foreign real short-term interest rate can vary only if another term in the condition—such as the expected change in the real exchange rate, or shocks to the UIP condition—also fluctuates (doing so in a way that makes the condition hold). In the flexible-price new open-economy model of Benigno and Thoenissen (2008), for example, there are no UIP shocks and few other features in the model that occasion the real exchange rate to vary much, so domestic and foreign real interest rates largely move in lockstep.

Does the fact that, under floating rates, real interest rates in a flexible-price small open-economy model are insensitive to the country's monetary policy—and are typically linked closely to real rates abroad—mean that monetary policy autonomy is absent from this kind of model? The answer is no. The linkage between real interest rates across countries is a sign of financial integration—not of the absence of monetary policy autonomy. And the central bank's inability in the home economy to affect real interest rates is a manifestation not of lack of monetary policy autonomy, but, instead, of the dichotomy between real and nominal variables that is a feature of flexible-price models. The central bank in the open economy has the power to set nominal variables even in the face of its inability to affect real variables. Indeed, in Benigno and Thoenissen's (2008) model, the central bank chooses and sets the inflation rate every period.[23]

To highlight the prevalence of results of the kind just described, it is useful to mention here two other recent contributions to the new open-economy literature. The first contribution worth highlighting is that of Benigno and Benigno (2008). These authors, for the most part, consider a sticky-price model. However, they are also concerned with behavior of the corresponding flexible-price economy. In particular, they consider the case of floating exchange rates and flexible prices. In this setting, it is feasible and appropriate for the monetary authority to set the inflation rate to zero every period (p. 982). Monetary autonomy—properly conceived—thus prevails under floating rates. At the same time, nominal interest rates enjoy a continuous relationship with rates abroad because of the uncovered interest parity condition (p. 973). The flexible-price version of the Benigno-Benigno model thus exhibits the simultaneous operation of the two quite distinct features of monetary autonomy and financial interdependence.[24]

A second study whose results underscore the coexistence of these two features is Devereux and Yetman (2010). A concern of that study is to model financial interdependence between countries in such a manner that cross-country financial linkages have nontrivial implications for the behavior of macroeconomic variables in each country. To this end, the authors embed in their model a balance sheet channel that "gives rise to a separate financial transmission mechanism of business cycle shocks that is independent of trade linkages" (Devereux and Yetman 2010, pp. 72–73). They go on to examine

[23] In their model, both the home economy and the foreign economy choose an inflation rate of zero. It would be possible, however, for the home economy to choose a different inflation rate from that prevailing abroad. Thanks to the Fisher effect, this would also imply a different nominal interest rate from that abroad, as expected inflation would differ across economies.

[24] As discussed below, the sticky-price version of the model also demonstrates the coexistence of these two properties.

model dynamics under a variety of specifications of financial market structure. Most of these specifications imply considerable financial interdependence among countries. Yet across specifications, the assumption is maintained of complete flexibility of nominal prices in each country and of floating exchange rates across countries. The model setting is consequently one in which inflation in each country is perfectly controllable by the country's central bank. Therefore, the various model settings in Devereux and Yetman (2010) all correspond to cases in which financial integration across countries coexists with monetary policy autonomy of the central bank in each country.

3.2 Floating Exchange Rates and Monetary Autonomy in Sticky-Price Models

Having considered these examples of flexible-price models, let us now turn to a situation in which a home economy has (temporarily) sticky prices. New open-economy models suggest that—depending on the precise specification of price stickiness and whether the stickiness applies to the entire price index or only to a subset of prices—the monetary authority may or may not have the ability in this environment to make the inflation rate equal to its target rate on a period-by-period basis. But irrespective of whether it can set the inflation rate every period in these sticky-price models, the central bank can set the long-run inflation rate equal to its target. Also irrespective of whether the central bank can set the inflation rate every period, it can set the nominal interest rate every period and, if desired, make it different from the rest-of-the-world nominal interest rate in any period. These generic features of sticky-price new open-economy models attest to the fact that the central bank has monetary policy autonomy under floating exchange rates in these models.[25] For example, in the sticky-price version of the Benigno and Benigno

[25]In Woodford's (2010) model, the combination of the assumed forms of nominal rigidity, international transactions, and the consumption bundle implies that the real interest rate (at all maturities) in the home economy is equal every period to the rest-of-world rate. This is so, even though in the corresponding *closed-economy* model with nominal rigidity the central bank *would* be able to influence the real interest rate in the short run. However, as discussed presently, in other new open-economy models the domestic real interest rate can be influenced by the small economy's central bank. And, as Woodford stresses, even in

(2008) model, floating rates allow each of the two countries in the model to follow rules for the short-term nominal interest rate that respond solely to their own country's inflation rate and output (see their proposition 2, p. 978).

Inflation-targeting central banks tend to emphasize their ability, via monetary policy actions, to make the short-term real interest rate deviate, in the short run, from the corresponding rate prevailing abroad. For example, Debelle (2017) clearly indicates that the RBA can influence Australian short-term real interest rates and so can make them behave differently from rest-of-world rates.[26] In the area of modeling, Romer (2000, p. 164) argues that it is vital for a realistic characterization of a small open economy that, under floating rates, the small open economy's central bank is able to generate differences between the path of the home short-term real interest rate and the path of the rest-of-world rate. Relatedly, Clarida (2019) and Holston, Laubach, and Williams (2017) provide empirical evidence that variation across countries in real policy interest rates might be warranted by country-specific shocks to the short-term natural real rate of interest.

In new open-economy models with floating rates, it is not a trivial matter to obtain settings in which the central bank can make movements in the real—and not just the nominal—short-term interest rate differ in the short run from those in the rest of the world. However, such settings are obtainable in certain variants of the new open-economy model of Erceg, Gust, and López-Salido (2010), for example. It is still the case in this environment that there is a *tendency* for the real interest rate to move in step with that abroad. But the central bank can, if it chooses, offset part or all of the influence of rest-of-world factors on the domestic real short-term interest rate by taking monetary policy actions that affect the domestic short-term interest rate (both nominal and real). The quid pro quo of

his model (which is based on that of Clarida, Galí, and Gertler 2002), the domestic central bank has monetary autonomy under floating exchange rates because it can always make the paths of the inflation rate, the nominal interest rate, and nominal aggregate demand depart from those prevailing abroad.

[26] Similarly, in a paper written shortly after his service as Governor of the Bank of England, Mervyn King suggested that central banks can make real interest rates in their country different from rest-of-world rates in the short run (see King and Low 2014, p. 3).

varying the home economy's short-term real interest rate in relation to the rest-of-world real rate is that the expected change in the real exchange rate will be nonzero.[27]

Now consider a special case of the above scenario: one in which prices are sticky not only at home but also in the rest of the world. Then the foreign central bank can affect foreign real interest rates. In addition, as already noted, there is a tendency, other things being equal, for real interest rates to move in step across countries. In combination, these model features imply that monetary policy actions in the rest of the world that affect the rest-of-world real short-term interest rate will tend to produce the same movement in the short-term real interest rate in the home economy. But such a situation does not mean that the central bank in the home economy lacks monetary policy autonomy. Foreign monetary policy is a force affecting the short-term real interest rate at home; but the central bank at home can itself exert an influence on the short-term real rate, possibly in a manner that offsets foreign influences on that rate.

The central bank therefore has monetary autonomy in sticky-price models with floating exchange rates. It can use monetary policy to set the long-run inflation rate and, thanks to sticky prices, is also able to pursue a stabilization goal for real output and can influence the short-term real interest rate. By choosing monetary policy's reactions to shocks, the central bank can shape the economy's overall short-run response to domestic and foreign real shocks and can make that response different from the response that would prevail in a fixed-exchange-rate regime.[28]

[27]This of course means that the UIP condition is satisfied.

[28]Banerjee, Devereux, and Lombardo (2016), for example, find that a monetary policy based on domestic goals is feasible under floating exchange rates (and full capital mobility) in their model, and they highlight one such policy that gives better economic outcomes than those prevailing under fixed exchange rates. Their conclusion does include the statement (p. 296) that "the benefits of flexible exchange rates and inflation targeting are very unlikely to hold in a global financial environment dominated by the currency and policy of a large financial center, such as [in] the current situation." This statement is an unwarranted concession to the no-autonomy position. The Banerjee-Devereux-Lombardo model results do affirm the existence of monetary autonomy under floating rates. The "benefits of flexible exchange rates" that the authors suggest do not hold in their model are features that are not part of the (properly stated) monetary-autonomy argument. (For example, the authors find that capital flows matter for economic

It may seem like too much emphasis has been given in the preceding paragraphs to the central bank's capacity to make short-term interest rates at home deviate from those abroad. After all, Rey (2016), while arguing against monetary autonomy, is willing to grant the ability of an open economy's central bank to affect short-term interest rates at home. But, in fact, the argument against autonomy is crucially undermined once it is established that central banks can set a country-specific course for short-term interest rates, as this is tantamount to establishing that monetary policy is not automatically and rigidly driven by external developments.

Rey (2016, pp. 24–25) takes as a violation of monetary autonomy her proposition that it is not possible, "relying on the domestic interest rate alone[,] to achieve both output stabilization and financial stability." Such a proposition is not, in fact, a contradiction of monetary autonomy. Monetary autonomy does not imply the possibility, under a float, of perfect stabilization of multiple goals using monetary policy. Instead, as emphasized earlier, monetary autonomy allows an open economy's monetary authority to select for their country a particular long-run inflation rate.[29] Beyond that basic decision, the monetary authority in a floating-rate economy would likely also use its autonomy to imbue monetary policy with other country-specific characteristics. The interest rate instrument can be deployed to achieve (or to trade off, in the event of a conflict) different macroeconomic-stability objectives—implying "the use of monetary policy for stabilization purposes," as Obstfeld and Rogoff (1995, p. 74) put it. The central bank can use monetary policy in pursuit of these macroeconomic stabilization objectives, and it can secure its desired long-run inflation rate, even in conditions in which monetary policy abroad is an influence on domestic economic activity.

Adding financial stability to the goals assigned to the country-specific monetary policy might well introduce a further tradeoff for monetary policy—in particular, between output stabilization and financial stability. But the existence of this tradeoff would not imply

activity in a floating-exchange-rate economy—a result not denied in the standard autonomy argument.)

[29] In practice, the selection would need to be one consistent with the central bank's statutory mandate.

lack of monetary autonomy. Indeed, even in a closed economy—in which the central bank invariably has monetary autonomy—a conflict could emerge between continuous achievement of output stabilization and a continuous condition of financial stability.

3.3 The Case in Which There Are Interest Rates Other Than the Policy Rate

In the sticky-price new open-economy models described above, the asset variables in the home economy that appear explicitly in the model equations are typically the exchange rate and a short-term security yield (which is used by the domestic monetary authority as a policy instrument). In practice, of course, an open economy has a broader spectrum of asset prices, among them equity prices and prices of longer-term securities. Some indication of how these asset prices might be expected to behave under a floating-exchange-rate regime is therefore in order. This is especially warranted in view of the fact that, as discussed in the next section, Rey (2016) sees the behavior of broader asset prices in advanced inflation-targeting economies as inconsistent with claims that the central banks of those economies possess monetary autonomy.

Let us first consider the situation in which there is no activity in longer-term securities markets by either the home or foreign central bank. The case of central bank purchases of longer-term securities will be considered subsequently.

In the environment laid out above—floating exchange rates, full capital mobility, and central banks' reliance on a short-term interest rate instrument—does evidence that international factors affect the domestic economy's asset prices, such as equity prices or prices of longer-term securities, constitute evidence against monetary autonomy? It would do so if the argument for autonomy claimed that, once a nation chooses to float its exchange rate, international factors, including capital flows, affect the exchange rate but have no effect on other asset prices in the home economy. However, the standard argument does not contain this claim. As stressed above, the standard concept of monetary autonomy instead involves the more modest claim that the home monetary authority can choose a course for its policy instruments—thereby influencing real conditions in the short run and nominal conditions in both the short run and the long run.

As already indicated, international capital mobility encourages co-movement of asset prices. It is in that light that Friedman and Friedman (1984, p. 127) noted that, under floating rates, a capital inflow puts downward pressure on domestic interest rates. What monetary autonomy provides is the opportunity for the central bank *also* to affect domestic interest rates. In particular, and as assumed here, a small country's central bank could undo the influence of international factors on the short-term interest rate and instead set a value for that rate, which consequently becomes the central bank's policy instrument. But for a given path of the current and expected policy rate, other asset prices in the small open economy will be function of world variables (inter alia), including rest-of-world monetary policy. For example, equity prices and the term-premium component of longer-term rates will likely have such a connection to world variables.

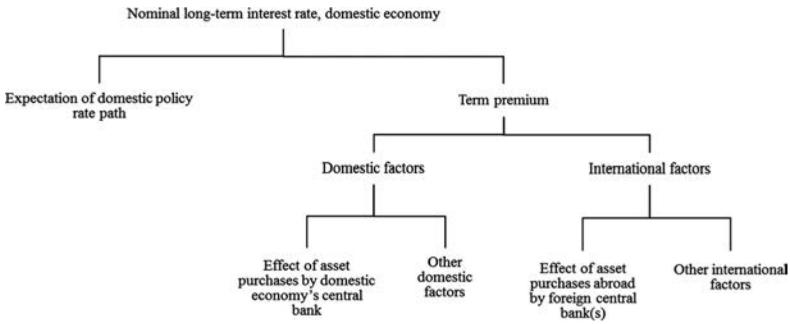
Long-Term Interest Rates and Asset Purchases. Although Rey (2016) considered the international effects of a conventional U.S monetary policy shock, the validity of the trilemma has also been discussed in the context of the Federal Reserve's large-scale asset purchases.[30] As is well known, these purchases, which took place principally from 2008 to 2014, were an unconventional monetary policy operation initiated by the Federal Open Market Committee (FOMC) and consisted of acquisitions of U.S. longer-term government securities, made with the intention of lowering the term premiums in the interest rates on these securities.[31] It is worth laying out how the notion of monetary autonomy in floating-rate small economies endures in the presence of asset purchases by a large economy's central bank. In what follows, it will be taken for granted that asset purchases are indeed capable of lowering term premiums.

Because of global trading in securities, the behavior of U.S. longer-term interest rates is likely one influence on longer-term interest rates in smaller countries. Therefore, in the presence of asset purchases, foreign central bank purchases likely become one of the

[30] See, for example, Taylor (2016, p. 242).

[31] Such a lowering was designed to provide stimulus to aggregate spending and so help bring U.S. economic outcomes closer to the FOMC's statutory objectives.

Figure 1. Influences on the Nominal Longer-Term Interest Rate in a Small Open Economy



factors affecting domestic as well as foreign term premiums. This is a way in which foreign monetary policy actions matter for the behavior of domestic long-term interest rates even when the exchange rate floats and short-term interest rates at home are set by the domestic monetary authority.

Does such an influence of foreign monetary policy on domestic longer-term interest rates imply a violation of monetary policy autonomy? The answer is no. For one thing, it remains the case under a float that monetary policy is not automatically and rigidly driven by external developments, and the central bank in the home economy can consequently choose a course for the short-term interest rate. But suppose—to take an extreme case—that the long-term interest rate in the domestic economy is the *only* interest rate or asset price that matters for aggregate demand in the open economy. In the presence of an influence of foreign monetary policy on this rate, it is still the case that there is monetary policy autonomy, because the home central bank is *also* able to influence the longer-term rate.

As usual, the longer-term interest rate can be decomposed into an expectation of the path of the short-term interest rate (which, by assumption, is the policy rate) and a term-premium component. See figure 1. The term premium is in turn affected by domestic and international factors. The international factors include the foreign central bank's asset purchases (that is, its purchases of securities issued by the government of its own country—purchases that produce a reaction of foreign term premiums, which in turn influence domestic term premiums). But the central bank in the home economy can itself affect longer-term interest rates at home through two means: by affecting the expected path of its policy rate and by its own asset purchase program. This central bank can therefore affect aggregate demand even when the long-term interest rate is the key rate for aggregate demand and when the long-term rate is partially determined by international factors.

4. Reconsideration of Empirical Evidence on Monetary Policy Autonomy

The discussion in the preceding sections implies that monetary autonomy is perfectly consistent with international financial integration. Monetary autonomy does not mean that there are no capital flows or that those flows do not affect domestic interest rates and asset prices. A central bank concerned with managing short-term interest rates can, through its market operations, enforce its will on a particular short-term market interest rate. The central bank thereby has a considerable influence on other domestic short-term interest rates. But both domestic and international forces will drive the overall constellation of asset prices.

It is thus evident that the case for floating exchange rates does not correspond to, or embed, a claim that a floating-rate country obtains *financial* independence from the rest of the world. On the contrary, provided that there is international capital mobility, one should observe financial integration across countries irrespective of the exchange rate regime. This bears very much on the validity of much of the case made against the feasibility of monetary policy autonomy under floating exchange rates. We will see that criticisms of autonomy have failed to take into account adequately, in the interpretation of evidence and in associated policy conclusions, the fundamental distinction between financial interdependence and monetary autonomy. In mischaracterizing what the monetary-autonomy argument claims about floating rates, the critique has invalidly taken international influences on domestic asset prices as evidence against monetary autonomy.

In an open economy with a floating exchange rate, it is to be expected that many asset prices will largely move in tandem with asset prices abroad and will be influenced by capital flows. Such financial integration does not preclude, or constitute evidence of the absence of, monetary policy autonomy for a small country whose exchange rate floats. The evidence that has been offered that open-economy central banks lack autonomy is, at bottom, based on the invalid premise that such autonomy implies a complete disconnection of asset-price movements across economies.

Nor does the argument for monetary autonomy suggest that the exchange rate is the only variable in the economy that adjusts to shocks from abroad. The exchange rate regime matters for how the economy responds to these international shocks. But a floating rate in itself does not insulate the economy's level of real output from the influence of international shocks in the short run, and it cannot prevent long-run adjustment of output to international shocks. Additionally, the cyclical behavior of nominal variables—like aggregate nominal spending and inflation—will likely be influenced by international shocks under a floating-rate regime. However, the central bank will be able to exercise a decisive influence on nominal variables over longer periods.

In light of these points, let us now consider the results of Rey (2016), which she sees as evidence against the existence of monetary autonomy in floating-rate advanced economies.

4.1 Rey's (2016) Evidence against Monetary Autonomy Reconsidered

As already stressed, a prominent means by which central banks are viewed as exercising monetary autonomy entails managing a short-term interest rate in their economies and making decisions on that rate in light of developments in domestic economic variables. However, in contending that inflation-targeting economies lack monetary policy autonomy, Rey (2016) does not focus on relationships between policy rates. She finds only mixed evidence of effects of U.S. monetary policy shocks on other advanced economies' policy rates and downplays the worth, as a metric for judging monetary policy autonomy, of cross-country co-movements of policy rates (see Rey 2016, pp. 7, 22, 24, 26).[32] Rather, she concentrates on reactions of domestic financial variables other than the policy rate in arguing that advanced economies lack monetary policy autonomy. Specifically, Rey (2016) conducts a vector autoregression (VAR) analysis of the effects of U.S. monetary policy shocks on key variables in several inflation-targeting advanced economies: the United Kingdom, Canada, Sweden, and New Zealand. In what she regards as a contrast with the notion that these nations' central banks can achieve domestic objectives under a float, she finds that financial variables other than the policy rate respond to a U.S. monetary shock.

In evaluating this evidence, it is worth beginning with one domestic variable that, although it is not actually included in her VAR analysis, Rey (2016, p. 10) highlights as affected by foreign factors (specifically, by U.S. monetary policy): credit growth. She stresses

[32]Rey is critical of other work on monetary autonomy that centers on the cross-country connections between policy rates. Indeed, her criticism includes the curious statement (Rey 2016, p. 24): "The trilemma, by focusing exclusively on the interest rate, seems to miss a potentially important channel of transmission of monetary policy in international markets." The reason this statement is curious is that, as already noted, the Mundell (1963) article cited by Rey as a key reference on the trilemma actually presented a model in which the monetary authority had *no* ability to influence interest rates under floating rates (a setting also described in Fleming 1962, p. 372). The Mundell (1963) reference is therefore itself sufficient to establish that the trilemma concept and the literature associated with it cannot be regarded as "focusing exclusively" on interest rates.

In a similar vein, Rey (2016, p. 7) specifically states that in the Mundell-Fleming model a floating exchange rate gives the central bank the power to control domestic interest rates. In fact, in the Mundell-Fleming model under full capital mobility, interest rates never differ from those abroad, irrespective of exchange rate regime. In early work by Mundell (such as Mundell 1960, 1961) a small open economy used the interest rate as its policy instrument, but his most celebrated contributions in the 1960s on the flexible-rate/fixed-rate distinction did not treat the interest rate as chosen by a small economy's central bank. (Indeed, in Mundell and Swoboda 1969, pp. 262–63, Mundell actually repudiated the practice of treating the interest rate as a policy instrument in an open economy.) It was not Mundell but other authors (including Friedman, as indicated above, and many contributors to the more recent new open-economy literature) who associated monetary autonomy with the ability of the central bank to influence the (or an) interest rate.

All this said, it is indeed the case that caution is needed in considering research that judges the existence or extent of monetary policy autonomy on the basis of cross-country correlations of policy rates. Such work has limitations discussed in section 4.2 below.

that credit growth in floating-rate open economies is influenced by international factors, operating via capital flows. However, it is almost a truism that a capital inflow will increase a country's credit growth irrespective of exchange rate regime, and this truism can hardly be regarded as evidence against monetary autonomy. A net capital inflow will tend to increase a country's overall external liabilities; and, usually, part of this increase in liabilities will take the form of borrowing from abroad. Therefore, ceteris paribus, capital inflows will tend to increase credit growth in the domestic economy. Under floating rates, the existence of this mechanism does not imply lack of monetary autonomy on the part of the home economy's central bank. On the contrary, under floating rates the capital inflow does not compromise the central bank's ability to manage short-term interest rates and, in particular, does not prevent the central bank from making short-term rates differ from those abroad.

Rey (2016, p. 13) nonetheless sees capital flows as violating monetary autonomy for the following reason: "As capital flows respond to U.S. monetary policy, they may not be appropriate for the cyclical conditions of many economies." But such a scenario actually has no bearing on the matter of monetary policy autonomy. An influence of capital flows on the domestic economy's business cycle is a case in which a development other than domestic monetary policy affects aggregate demand. It does not imply that the central bank is not itself incapable of affecting (real and nominal) aggregate demand. Provided that the central bank has this capability, it can, if it desires, take actions that offset other forces, including capital flows, affecting aggregate demand. That is, under floating exchange rates, the central bank is able to set the short-term nominal interest rate in response to the economic outlook, including any effect that capital flows have had on the outlook. Furthermore, and in contrast to the fixed-rate case, the central bank under floating rates is able to decide, via its monetary policy response, whether capital flows are permitted to have any lasting influence on nominal variables.

Let us now consider the key domestic financial variable (other than the policy rate) included in Rey's (2016) VAR analysis.[33]

[33]Rey (2016) also shows responses to U.S. monetary policy shocks of inflation-targeting countries' output and consumer price index. It is consistent with monetary autonomy for output to be responsive to international shocks (including,

This is the "mortgage spread"—defined as the spread between the mortgage interest rate and the long-term government bond rate.

Central banks in open economies certainly tend to take the position that the mortgage interest rate is an important interest rate that their own actions influence. But Rey's results do not actually contradict this position. She finds that U.S. monetary policy shocks are a factor affecting the spread in open economies between the mortgage interest rate and the long-term government bond rate. Nothing in this result precludes the home central bank's monetary policy from being an important factor affecting the *absolute level* of mortgage interest rates.

With regard to long-term government bond rates themselves, these—like equity prices—fall into the category of domestic asset prices that one would expect to be influenced by world conditions, including U.S. monetary policy shocks, even when the exchange rate floats and the central bank has monetary policy autonomy. This point was discussed in section 3.3 above. Nothing in Rey's results is inconsistent with the notion that longer-term rates are *also* affected by domestic factors, including the actions of the home central bank.

In sum, Rey's (2016) empirical findings are consistent with financial integration of the kind that one should observe when a small open economy possesses the features of international capital mobility, floating exchange rates, and monetary policy autonomy. It follows that results indicating that movements in asset prices in the economy are linked to those in the rest of the world do not constitute evidence against that economy's monetary policy autonomy.

Rey (2016, p. 13) states: "Although seeing a lot of comovement in asset prices worldwide may just be reflecting market integration, the fact that these comovements are to some extent caused by U.S. monetary policy is important." Such co-movements and their source may indeed be important. But these factors do not refute the existence of monetary policy autonomy. Their existence does not obstruct the home central bank from affecting the actual real policy rate.

in the short run, foreign monetary policy shocks that affect real asset prices and output in the foreign economy). And while, under monetary autonomy, it is possible for monetary policy to make nominal variables such as the consumer price index (CPI) immune from foreign influences over long periods, autonomy does not imply that the CPI and CPI inflation are insensitive to international shocks in the short run.

The central bank is then in a position to reinforce, accommodate, or offset international shocks that affect the domestic economy's level of aggregate demand. In Wicksellian terms, international shocks are among the factors determining the natural value of the real policy rate in the domestic economy: such shocks affect the value of the actual policy rate consistent with keeping output at potential. But their existence does not prevent the home economy's monetary policy, as reflected in the central bank's choice for the policy rate, from influencing aggregate demand.[34]

4.2 Interrelations of Policy Rates across Countries

Let us now return to the situation in which monetary policy is centered on management of a short-term interest rate. A great deal of research on testing and indexing monetary policy autonomy focuses on the strength of the correlations of short-term interest rates across countries under different exchange rate regimes.[35] For example, Edwards (2017, p. 10), like many others, notes that lack of monetary policy autonomy implies a situation in which a small country's policy rate must keep in step with the short-term nominal interest rate prevailing in the rest of the world, and his tests of monetary autonomy rest on the premise that autonomy is associated with statistical independence of the home policy rate and the rest-of-world rate. The same notion underlies such investigations as Aizenman, Chinn, and Ito (2008) and Klein and Shambaugh (2015).

There are, however, grounds for believing that, outside the extreme cases of correlations of zero or unity, positive correlations of policy rates across countries are not informative about the degree of monetary policy autonomy. To be sure, a perfect or near-perfect positive correlation of a small country's policy rate with the policy rate abroad is very likely testament to fixed-exchange-rate conditions (formally or de facto) and to a corresponding lack of monetary autonomy. And the ability of an open economy's central bank to make its policy rate wholly uncorrelated over long periods with that

[34] This is consistent with Debelle's (2017) account, quoted in section 2.4, of the situation facing the Australian monetary authorities.

[35] For an early study in this vein, see Throop (1980).

in economies abroad does point to the likelihood that it possesses monetary autonomy.

In contrast, the intermediate case of an imperfect positive correlation is unlikely to be informative about autonomy. Findings that, under capital mobility and floating rates, policy rates are correlated across countries, or results indicating that the domestic policy rate responds to foreign variables, do not in themselves amount to a refutation of the existence, availability, and practical importance of monetary policy autonomy in small open economies whose exchange rates float.

To see this, it is useful to put aside the case in which the small open economy's central bank decides to follow a monetary policy strategy not solely oriented toward domestic aggregate goals. This case includes the possibilities that the central bank implements a managed exchange-rate float or responds to developments in the trade sector—as Bernanke (2017) argues might be the situation in some emerging economies. Such possibilities will be bypassed, as they are unlikely to pertain to the advanced economies—focused upon here—that target inflation and float their exchange rate. The monetary authorities in these latter economies likely instead fit Debelle's (2017) characterization of being concerned with "very much domestic" objectives.

In the case of these domestic-objectives-focused economies, a positive correlation between policy rates and those abroad might emerge under monetary autonomy for two basic reasons. First, there could be policy rate responses to international shocks that matter for—and might tend to destabilize, absent an appropriate policy response—domestic variables like inflation or the output gap.[36] Second, a positive correlation could arise from the central bank's response to shocks to domestic spending and production that have an international component, in the sense of being correlated with the corresponding shocks elsewhere. For example, taste or technology shocks might have a global component.

[36]Such a domestic policy response to international shocks might be brought about via the policymakers' belief in the importance of the exchange rate for the behavior of domestic variables. For example, King (1997, p. 227) observed that in U.K. inflation targeting, exchange rate behavior was "an important component of our assessment of the economy and the prospects for inflation."

The point that a policy strategy focused on domestic economic stability might involve responses to global shocks has occasionally been acknowledged, but not stressed, in studies that focus on policy rate correlations in judging the validity of the trilemma (see, for example, Klein and Shambaugh 2015, p. 38). It receives more emphasis in Obstfeld (2015, p. 14). However, the shocks possessing a global component likely go beyond the financial shocks that Obstfeld nominates as candidates. Global financial shocks certainly are a plausible source of cross-country macroeconomic interaction—especially if cross-country financial channels of the kind modeled by Devereux and Yetman (2010) are important—and could lead to correlations across countries of policy rates; but, as already indicated, shocks to spending on goods and to the supply side also likely have an important global component.

The reasoning outlined above implies that, even though monetary autonomy can make an open economy's policy rate statistically independent of policy rates abroad, a strictly positive, but imperfect, correlation with policy rates abroad need not signify the absence of monetary autonomy.

5. Conclusion

In a closed-economy model, the monetary authority is able to exert a decisive influence on the course of nominal variables in the long run and, if there is price stickiness that wears off over time, it is also able to influence the course of real variables in the short run. If a small economy's central bank in an open-economy model has these same abilities when it lets its exchange rate float, then it is appropriate to conclude that the central bank in that model has monetary policy autonomy under a float. It has been argued above that the central bank does have such autonomy under floating rates in standard new open-economy models and that, furthermore, this autonomy is of practical relevance, aiding understanding of policy behavior and economic outcomes in floating-rate, inflation-targeting advanced economies.

Monetary autonomy for a central bank, as expounded by Friedman (1953a) in his case for floating exchange rates, means that the central bank has prerogatives regarding the creation of the total amount of base money, irrespective of the monetary policy pursued abroad. A corollary of this is that, for a small open economy whose exchange rate floats, variations in trade balances, in capital flows, or in monetary developments abroad have no automatic implications for the setting of monetary policy instruments. Consequently, the economy's central bank is in a position to pursue stabilization policies (typically, in practice, via the management of a short-term interest rate). That situation holds under floating exchange rates in actuality, as well as in sticky-price new open-economy models under the assumption of floating.

Monetary autonomy does not imply that asset prices at home (including interest rates other than the policy rate) are insensitive to international factors, including developments in monetary policy abroad. On the contrary, everything else equal, financial integration will create tendencies for real yields in the home economy to move closely with those in the rest of the world. When prices are sticky in the rest of the world, foreign monetary policy will be one influence, in the short run, on rest-of-world real yields and so on domestic real yields, without any violation of monetary autonomy in the home economy.

In light of these properties and implications of monetary autonomy, it is clear that empirical evidence like that recently offered as evidence against the practical importance of monetary autonomy does not, in fact, amount to valid evidence against autonomy. As this paper has stressed, empirical findings put forward as contradicting monetary autonomy—such as the influence of foreign monetary policy shocks on domestic asset prices—are, in qualitative terms, features that can be found in new open-economy models in which the monetary authority possesses autonomy under floating exchange rates and complete capital mobility.

Rey (2016) finds that U.S. monetary policy shocks affect asset prices and other financial conditions in advanced inflation-targeting economies. Such evidence confirms that foreign monetary policy likely is one of the sources of international shocks that affect output and aggregate demand in open economies. But, as indicated above, monetary policy autonomy does not require that the domestic economy is unaffected by shocks abroad, including, in the short run, foreign monetary policy shocks. It only requires that the central bank at home is itself able to affect the domestic economy by influencing aggregate demand.

Edwards (2017) and others find that policy rates in small open economies are related to those abroad, even under conditions of floating exchange rates. But, as has also been indicated above, although a float enables the central bank to create deviations in the policy rate from policy rates abroad, its pursuit of domestic objectives may lead it, on occasion, to make the policy rate move with rest-of-world rates. Consequently, although a lack of correlation between the home policy rate and that abroad may be testament to monetary autonomy in the home economy, a significant positive correlation can be consistent with the operation of an autonomous monetary policy, directed at domestic objectives, in that economy.

The analysis given here reaffirms the standard result that if an open economy floats its exchange rate, it secures monetary policy autonomy. Controls on international capital movements are not needed for autonomy. The result that monetary autonomy prevails whether capital controls are imposed or not does not, of course, mean that financial conditions in the home economy are the same with or without such controls. On the contrary, capital controls modify—and likely reduce—the influence of rest-of-world developments on asset prices and on other financial conditions in the home economy. It is possible that the authorities of an open economy could see merit, on net, in such a situation—perhaps on financial stability grounds.[37] Nevertheless, such considerations do not bear on the validity of the dilemma/trilemma distinction—which pertains to monetary policy's power. The linkage between the economy's asset prices and those in the rest of the world should be clearly distinguished from the central bank's ability to carry out a stabilization policy that shapes the path of aggregate demand and inflation. Provided that it has that power in a floating-rate setting (even without capital controls),

[37]It is tempting to add that, as asset prices are among the factors that influence aggregate demand, capital controls imply greater scope for the monetary authority in its demand management. But it bears repeating that monetary autonomy refers to monetary policy's ability to affect aggregate demand; provided that policy has this ability under a float, monetary autonomy prevails. Capital controls might mean that there are fewer forces that monetary policy needs to offset to stabilize aggregate demand; the monetary authority may feel that the task of stabilization policy is thereby made easier. But monetary autonomy would prevail even without imposition of these controls.

a central bank in an open economy does possess the monetary policy autonomy described by the trilemma.

References

Aizenman, J., M. D. Chinn, and H. Ito. 2008. "Assessing the Emerging Global Financial Architecture: Measuring the Trilemma's Configurations Over Time." NBER Working Paper No. 14533 (December).

Banerjee, R., M. B. Devereux, and G. Lombardo. 2016. "Self-Oriented Monetary Policy, Global Financial Markets and Excess Volatility of International Capital Flows." *Journal of International Money and Finance* 68 (1, November): 275–97.

Beaumont, C., and L. Cui. 2007. "Conquering Fear of Floating: Australia's Successful Adaptation to a Flexible Exchange Rate." IMF Policy Discussion Paper No. 07/2 (July).

Benigno, G., and P. Benigno. 2008. "Exchange Rate Determination under Interest Rate Rules." *Journal of International Money and Finance* 27 (6, October): 971–93.

Benigno, G., and C. Thoenissen. 2008. "Consumption and Real Exchange Rates with Incomplete Markets and Non-Traded Goods." *Journal of International Money and Finance* 27 (6, October): 926–48.

Bernanke, B. S. 2017. "Federal Reserve Policy in an International Context." *IMF Economic Review* 65 (1, April): 5–36.

Bordo, M. D., and J. B. Taylor, eds. 2017. *Rules for International Monetary Stability: Past, Present, and Future*. Stanford, CA: Hoover Institution Press.

Clarida, R. 2019. "The Global Factor in Neutral Policy Rates: Some Implications for Exchange Rates, Monetary Policy, and Policy Coordination." *International Finance* 22 (1, Spring): 2–19.

Clarida, R., J. Galí, and M. Gertler. 2002. "A Simple Framework for International Monetary Policy Analysis." *Journal of Monetary Economics* 49 (5, July): 879–904.

Debelle, G. 2017. "Global Influences on Domestic Monetary Policy." Speech given to the Committee for Economic Development of Australia (CEDA), Adelaide, July 21.

Dellas, H., and G. S. Tavlas. 2018. "Milton Friedman and the Case for Flexible Exchange Rates and Monetary Rules." *Cato Journal* 38 (2, Spring/Summer): 361–77.

Devereux, M. B., and J. Yetman. 2010. "Leverage Constraints and the International Transmission of Shocks." *Journal of Money, Credit and Banking* 42 (s1, September): 71–105.

Dornbusch, R. 1986. *Dollars, Debts, and Deficits.* Cambridge, MA: MIT Press.

The Economist. 1992. "About-Face." September 26, p. 30.

Edwards, S. 2017. "Monetary Policy Independence under Flexible Exchange Rates: The Federal Reserve and Monetary Policy in Latin America—Is There Policy 'Spillover?'" In *Rules for International Monetary Stability: Past, Present, and Future,* ed. M. D. Bordo and J. B. Taylor, 1–41 (chapter 1). Stanford, CA: Hoover Institution Press.

Erceg, C. J., C. Gust, and D. López-Salido. 2010. "The Transmission of Domestic Shocks in Open Economies." In *International Dimensions of Monetary Policy,* ed. J. Galí and M. Gertler, 89–148 (chapter 2). Chicago: University of Chicago Press.

Federal Reserve Bank of Kansas City. 2013. "General Discussion: 'Dilemma Not Trilemma: The Global Financial Cycle and Monetary Policy Independence'." In *Global Dimensions of Unconventional Monetary Policy.* Kansas City, MO: Federal Reserve Bank of Kansas City.

Fleming, J. M. 1962. "Domestic Financial Policies under Fixed and under Floating Exchange Rates." *IMF Staff Papers* 9 (3, November): 369–80.

Friedman, M. 1953a. "The Case for Flexible Exchange Rates." In *Essays in Positive Economics*, ed. M. Friedman, 157–203. Chicago: University of Chicago Press.

———. 1953b. "Why the Dollar Shortage?" *The Freeman* 4 (6): 201–3. https://mises.org/library/freeman-december-1953. Reprinted with additions in Milton Friedman, *Dollars and Deficits: Inflation, Monetary Policy and the Balance of Payments.* Englewood Cliffs, NJ: Prentice-Hall.

———. 1975. "Six Fallacies." *Wall Street Journal*, June 30, 11.

———. 1983. "The Keynes Centenary: A Monetarist Reflects." *The Economist* June 4, 17–19.

Friedman, M., and R. D. Friedman. 1984. *Tyranny of the Status Quo*. New York: Harcourt Brace Jovanovich.

Friedman, M., and R. V. Roosa. 1967. *The Balance of Payments: Free Versus Fixed Exchange Rates*. Washington, DC: American Enterprise Institute for Public Policy Research.

Holston, K., T. Laubach, and J. C. Williams. 2017. “Measuring the Natural Rate of Interest: International Trends and Determinants.” *Journal of International Economics* 108 (S1, May): S59–S75.

Irwin, D. A. 2012. *Trade Policy Disaster: Lessons from the 1930s.* Cambridge, MA: MIT Press.

———. 2017. “The Missing Bretton Woods Debate over Flexible Exchange Rates.” NBER Working Paper No. 23037 (January).

Johnson, H. G. 1969. “The Case for Flexible Exchange Rates, 1969.” *Review* (Federal Reserve Bank of St. Louis) 51 (6, June): 12–24.

Keister, T. 2012. “Corridors and Floors in Monetary Policy.” Liberty Street Economics blog, Federal Reserve Bank of New York, April 4.

King, M. 1994. “Monetary Policy Instruments: The U.K. Experience.” *Quarterly Bulletin* (Bank of England) 34 (3, September): 268–76.

———. 1997. “Monetary Policy and the Exchange Rate.” *Quarterly Bulletin* (Bank of England) 37 (2, June): 225–27.

King, M., and D. Low. 2014. “Measuring the ‘World’ Real Interest Rate.” NBER Working Paper No. 19887 (February).

Klein, M. W., and J. C. Shambaugh. 2015. “Rounding the Corners of the Policy Trilemma: Sources of Monetary Policy Autonomy.” *American Economic Journal: Macroeconomics* 7 (4): 33–66.

Krugman, P. 1993. “Exchange Rates.” In *Fortune Encyclopedia of Economics*, ed. D. R. Henderson, 518–22. New York: Warner Books.

McCallum, B. T. 1996. *International Monetary Economics.* New York: Oxford University Press.

Meade, J. 1961. “The Future of International Trade and Payments.” *Three Banks Review* 13 (50, June): 15–38.

Mundell, R. A. 1960. “The Monetary Dynamics of International Adjustment under Fixed and Flexible Exchange Rates.” *Quarterly Journal of Economics* 74 (2, May): 227–57.

———. 1961. “Flexible Exchange Rates and Employment Policy.” *Canadian Journal of Economics and Political Science* 27 (4, November): 509–17.

———. 1963. “Capital Mobility and Stabilization Policy under Fixed and Flexible Exchange Rates.” *Canadian Journal of Economics and Political Science* 29 (4, November): 475–85.

———. 1968. “A Plan for a World Currency.” In *Next Steps in International Monetary Reform: Hearing Before the Subcommittee on International Exchange and Payments of the Joint Economic Committee, Congress of the United States, Ninetieth Congress, Second Session, September 9, 1968*, Joint Economic Committee, U.S. Congress, 14–28. Washington, DC: U.S. Government Printing Office.

Mundell, R. A., and A. K. Swoboda, eds. 1969. *Monetary Problems of the International Economy.* Chicago: University of Chicago Press.

Nelson, E. 2009. “Milton Friedman and U.K. Economic Policy, 1938–1979.” Working Paper No. 2009-017A, Federal Reserve Bank of St. Louis (April).

Obstfeld, M. 1998. “The Global Capital Market: Benefactor or Menace?” *Journal of Economic Perspectives* 12 (4, Autumn): 9–30.

——— 2015. “Trilemmas and Trade-offs: Living with Financial Globalisation.” BIS Working Paper No. 480.

Obstfeld, M., J. D. Ostry, and M. S. Qureshi. 2017. “A Tie That Binds: Revisiting the Trilemma in Emerging Market Economies.” IMF Working Paper No. 17/130 (June).

Obstfeld, M., and K. Rogoff. 1995. “The Mirage of Fixed Exchange Rates.” *Journal of Economic Perspectives* 9 (4, Fall): 73–96.

Obstfeld, M., and A. M. Taylor. 1998. “The Great Depression as a Watershed: International Capital Mobility Over the Long Run.” In *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century,* ed. M. D. Bordo, C. D. Goldin, and E. N. White, 353–402 (chapter 10). Chicago: University of Chicago Press.

———. 2017. “International Monetary Relations: Taking Finance Seriously.” *Journal of Economic Perspectives* 31 (3, Summer): 3–28.

Pasricha, G. K. 2017. “Policy Rules for Capital Controls.” BIS Working Paper No. 670 (November).

Rey, H. 2013. "Dilemma Not Trilemma: The Global Financial Cycle and Monetary Policy Independence." In *Global Dimensions of Unconventional Monetary Policy*. Kansas City, MO: Federal Reserve Bank of Kansas City.

———. 2016. "International Channels of Transmission of Monetary Policy and the Mundellian Trilemma." *IMF Economic Review* 64 (1, May): 6–35.

Romer, D. 2000. "Keynesian Macroeconomics without the LM Curve." *Journal of Economic Perspectives* 14 (2, Spring): 149–69.

Schmitt-Grohé, S., and M. Uribe. 2003. "Closing Small Open Economy Models." *Journal of International Economics* 61 (1, October): 163–85.

Taylor, J. B. 2016. "Rethinking the International Monetary System." *Cato Journal* 36 (2, Spring/Summer): 239–50.

Throop, A. W. 1980. "Managed Floating and the Independence of Interest Rates." *Economic Review* (Federal Reserve Bank of San Francisco) 6 (3, Summer): 6–23.

Woodford, M. 2001. "Monetary Policy in the Information Economy." In *Economic Policy for the Information Economy*, 297–370. Kansas City, MO: Reserve Bank of Kansas City.

———. 2010. "Globalization and Monetary Control." In *International Dimensions of Monetary Policy,* ed. J. Galí and M. J. Gertler, 13–77. Chicago: University of Chicago Press.

Output Gaps and Robust Monetary Policy Rules*

Roberto M. Billi
Sveriges Riksbank

Policymakers often use the output gap to guide monetary policy, even though inflation and nominal gross domestic product (NGDP) are measured more accurately in real time. Employing a small New Keynesian model with a zero lower bound (ZLB) on nominal interest rates, this article compares the performance of monetary policy rules that are robust to persistent measurement errors. It shows that, in the absence of the ZLB, the central bank should focus on stabilizing inflation rather than nominal GDP. But when the ZLB is present, a policy that seeks to stabilize nominal GDP improves substantially the tradeoffs faced by the central bank.

JEL Codes: E31, E52, E58.

*I thank for valuable comments Yunus Aksoy, Pierpaolo Benigno (coeditor), Fabio Canova, Larry Christiano, Emmanuel De Veirman, Andy Filardo, Jordi Galí, Tor Jacobson, Per Jansson, Keith Kuester, Stefano Neri, Damjan Pfajfar, Charles Plosser, Ricardo Reis, Frank Smets, Ulf Söderström, Lars Svensson, John Taylor, Robert Tetlow, David Vestin, Karl Walentin, John Williams, Mike Woodford, Andreas Wörgötter, seminar participants at Banca d'Italia, Bank of England, Copenhagen Business School, Federal Reserve Bank of Philadelphia, Joint Seminar Series of the ECB-CFS-Bundesbank, Sveriges Riksbank, Tilburg University, Uppsala University, as well as conference participants at AEA, Joint French Macro Workshop at Banque de France, Macro Workshop at De Nederlandsche Bank, Midwest Macroeconomics Meeting, Monetary Policy Challenges from a Small Country Perspective at National Bank of Slovakia, and SCE. I gratefully acknowledge the ECB's hospitality while working on this project in 2018. The views expressed herein are solely the responsibility of the author and should not be interpreted as reflecting the views of Sveriges Riksbank. Address correspondence to: Roberto M. Billi, Monetary Policy Department, Sveriges Riksbank, SE-103 37 Stockholm, Sweden; e-mail: Roberto.Billi@riksbank.se.

1. Introduction

In monetary policy analysis, a commonly used measure of economic activity is the output gap, which is a gauge of how far the economy is from its productive potential. The output gap is conceptually appealing as an indicator to help guide policy because it is an important determinant of inflation developments. A positive output gap implies an overheating economy and upward pressure on inflation. By contrast, a negative output gap implies a slack economy and downward pressure on inflation. Thus, if available, accurate and timely estimates of the output gap can play a central role in the conduct of effective monetary policy. A positive output gap prompts the central bank to cool an overheating economy by raising policy rates, whereas a negative output gap prompts for adding monetary stimulus.

In practice, however, the output gap is a noisy signal of economic activity. Estimates of the output gap are often subject to large revisions, even long after the time policy is actually made.[1] Thus, there is broad interest in finding monetary policies that are robust to persistent errors in measuring the output gap. As Taylor and Williams (2010) explained, one view is that in simple policy rules the optimal coefficient on the output gap declines in the presence of noise in measuring the gap. The logic for this result is straightforward. The reaction to the mismeasured output gap adds unwanted noise to the setting of monetary policy, which causes unnecessary fluctuations in output and inflation. Such adverse effects of noise can be reduced by lowering the coefficient on the output gap in the policy rule.

At the same time, many argue for greater policy activism when the zero lower bound (ZLB) on nominal interest rates constrains policy.[2] The inability to reduce the policy interest rate below its

[1]Measuring the output gap involves two complications. First, potential output cannot be measured directly, so it must be estimated. Second, GDP data are regularly revised as statistical agencies incorporate more complete source information and new methodologies into the published data. As Orphanides and van Norden (2002) showed, estimating potential output is the main source of errors in measuring the output gap.

[2]This article adopts the standard practice of referring to a zero lower bound on nominal interest rates. The recent experience with negative nominal interest rates

effective lower bound can limit, or even impair, the ability of monetary policy to stabilize output and inflation. As Reifschneider and Williams (2000) showed, increasing the coefficient on the output gap in simple policy rules can improve economic performance. An active response to the output gap prescribes greater monetary stimulus before and after episodes when the ZLB constrains policy, which lessens deflationary pressures when the ZLB constrains policy. But there are clear limits to such an approach, as it generally increases the volatility of inflation and interest rates. A large coefficient on the output gap can be counterproductive, especially when the output gap is mismeasured.

In light of such concerns, another perspective is that central banks should ignore the output gap altogether to focus strictly on stabilizing inflation or seek instead to stabilize the level of nominal gross domestic product (NGDP). NGDP-level targeting is particularly appealing for two reasons. First, monetary policy is then expected to be more robust to errors in the measurement of economic activity, because revisions to GDP are typically smaller than revisions to the output gap. Estimates of GDP are not prone to errors from estimating potential output. Second, the central bank is also required to make up for any past shortfalls from its NGDP target, which ensures greater policy stimulus during ZLB episodes.

This article, thus, studies the performance of such monetary policy rules in a small New Keynesian model, with the central bank facing persistent errors in the measurement of economic conditions and a ZLB constraint.[3] In the model, several types of structural and noise shocks buffet the economy. On the supply side, technology shocks push output gaps and inflation in the same direction, whereas cost-push shocks instead cause an inflation-output tradeoff. On the demand side, adverse demand shocks and the ZLB constraint create a tradeoff between stabilizing current and future output, because it is desirable for the central bank in a ZLB episode to promise to

in Denmark, Japan, Sweden, Switzerland, and the euro zone suggests the effective lower bound is somewhat below zero. See Svensson (2010) for a discussion.

[3]This analysis assumes the private sector possesses full information about the state of the economy in real time, which implies the model can be treated as structurally invariant under different policies. See Aoki (2003, 2006) and Svensson and Woodford (2004) for a discussion.

induce an expansion after the ZLB episode.[4] Moreover, the central bank faces persistent noise shocks in the setting of monetary policy, which creates a tradeoff between fluctuations in the economy from structural shocks and those from noise shocks.

The stylized model offers a clear illustration of such tradeoffs in the evaluation of the monetary policy rules. Before proceeding to the evaluation, the model is calibrated to recent U.S. data, with the conduct of monetary policy described by a simple rule often used in policy analysis, namely a version of the Taylor rule with interest rate smoothing. In the calibration of the model, the structural shocks are persistent to generate propagation in the model as in the data. The noise shocks are persistent to reflect historical revisions of the data. Also considered is the optimal commitment policy, to be used as a benchmark for the evaluation. The monetary policy rules are then ranked in terms of performance, based on the model's social welfare function.

With the calibrated model, I study the extent to which persistent errors in the measurement of economic conditions and a ZLB constraint adversely affect the performance of the two targeting rules and inertial Taylor rule, relative to the optimal commitment policy. The analysis produces three main results. First, under the optimal commitment policy, although measurement error and the ZLB constraint are both a source of fluctuations in output and inflation, social welfare is more severely affected by the ZLB constraint. As a second main result, the ZLB constraint plays a critical role for the ranking of two targeting rules. In the absence of the ZLB, the central bank should focus on stabilizing inflation rather than nominal GDP. But present the ZLB, a policy that seeks to stabilize the level of nominal GDP improves substantially the tradeoffs faced by the central bank. And third, if monetary policy becomes more severely constrained by the ZLB, social welfare is more severely affected under a targeting rule that does not require the central bank to make up for past shortfalls from the target.

In the previous literature, in the aftermath of the financial crisis and Great Recession, proposals for NGDP-level targeting include

[4]The promise is credible if the central bank commits to making up for past shortfalls from its target, as is the case with an inertial Taylor rule or with a NGDP-level target.

Hatzius and Stehn (2011, 2013), Sumner (2011, 2014), Woodford (2012, 2013), Frankel (2013), and Billi (2017), among others.[5] But none of these articles takes into account that central banks face persistent errors in the measurement of economic conditions. In another strand of literature, studies about the design of monetary policies that are robust to measurement error include Orphanides et al. (2000), Orphanides (2001, 2003), Rudebusch (2002), Smets (2002), Aoki (2003, 2006), Ehrmann and Smets (2003), Svensson and Woodford (2003, 2004), Boehm and House (2014), Garín, Lester, and Sims (2016), and others. But these articles do not take into account a ZLB constraint. Gust, Johannsen, and López-Salido (2017) study the interaction between mismeasurement of the state of the economy and a ZLB constraint but for tractability need to assume the mismeasurement is not persistent. Relative to the previous literature, the contribution of this article is to show that the ranking of the monetary policy rules depends crucially on the likelihood of hitting the ZLB constraint.

The article proceeds as follows. Section 2 describes the model and monetary policy rules. Section 3 presents the model outcomes and policy evaluation. Section 4 concludes. The appendix contains technical details of the model solution and additional results.

2. The Model

I use a small New Keynesian model as described in Woodford (2010). I describe the conduct of monetary policy with targeting rules, and with a simple rule to be used for the calibration of the model. In each of the policy frameworks considered, the central bank faces persistent errors in the measurement of economic conditions and a ZLB constraint. I explain the features of this model and the equilibrium, and then calibrate the model to U.S. data.

[5]There is also an extensive literature on the notion of nominal income *growth* targeting, at first suggested by Meade (1978) and Tobin (1980) and then studied by Bean (1983), Taylor (1985), West (1986), McCallum (1988), Clark (1994), Hall and Mankiw (1994), Jensen (2002), Walsh (2003), and Billi (2011b), among others.

2.1 Private Sector

The behavior of the private sector is described by two structural equations, log-linearized around zero inflation, which represent the demand and supply sides of the economy. The economy is buffeted by persistent demand and supply shocks.

On the demand side of the economy, the Euler equation describes the representative household's expenditure decisions,

$$y_t = E_t y_{t+1} - \varphi \left(i_t - r - E_t \pi_{t+1} - v_t \right), \tag{1}$$

where E_t denotes the expectations operator conditional on information available at time t. y_t is output measured as the log-deviation from a trend. π_t is the inflation rate, the log-change of prices from the previous period,

$$\pi_t \equiv p_t - p_{t-1}. \tag{2}$$

And $i_t \geq 0$ is the short-term nominal interest rate, which is the instrument of monetary policy and is constrained by a ZLB. $r > 0$ is the steady-state interest rate.[6] $\varphi > 0$ is the interest elasticity of real aggregate demand, capturing intertemporal substitution in household spending. The *demand shock*, v_t, represents other spending, such as government spending, which has asymmetric effects on the economy due to the ZLB constraint. A positive demand shock can be countered entirely by raising the nominal interest rate, whereas a large adverse shock that leads to hitting the ZLB causes an economic downturn.

On the supply side of the economy, the Phillips curve describes the optimal price-setting behavior of firms, under staggered price changes à la Calvo,

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t + u_t, \tag{3}$$

where $\beta \in (0, 1)$ is the discount factor of the representative household, determined as $1/(1 + r)$. The slope parameter $\kappa > 0$ is a function of the structure of the economy.[7] $x_t \equiv y_t - y_t^n$ is the output gap

---

[6]Thus, $i_t - r - E_t \pi_{t+1}$ is the real interest rate in deviation from steady state.

[7]In this model $\kappa = (1 - \alpha)(1 - \alpha\beta)\alpha^{-1}\left(\varphi^{-1} + \omega\right)(1 + \omega\theta)^{-1}$, where $\omega > 0$ denotes the elasticity of a firm's real marginal cost. $\theta > 1$ is the price elasticity of demand substitution with firms in monopolistic competition, and thus the

in the economy. y_t^n is the natural rate of output, or potential output, the output deviation from the trend that would prevail in the absence of any price rigidities, which represents a *technology shock*. A positive technology shock implies slack in economic activity and downward pressure on prices, whereas a negative shock implies a strong economy and puts upward pressure on prices. Moreover, u_t is a cost-push shock, or *markup shock* resulting from variation over time in the degree of monopolistic competition between firms, which creates an inflation-output tradeoff for monetary policy.

In this model economy, the three types of exogenous structural shocks (y_t^n, u_t, v_t) are assumed to follow AR(1) stochastic processes, with first-order autocorrelation parameters $\rho_j \in [0,1)$ for $j = y^n, u, v$. Moreover, $\sigma_{\varepsilon j}\varepsilon_{jt}$ are the innovations that buffet the economy, which are independent across time and cross-sectionally, and are normally distributed with mean zero and standard deviations $\sigma_{\varepsilon j} > 0$.

Finally, the monetary policy rules to be considered are evaluated based on the model's social welfare function, a second-order approximation around zero inflation of the lifetime utility function of the representative household,

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\pi_t^2 + \lambda\left(x_t - x^*\right)^2\right], \tag{4}$$

where $\lambda = \kappa/\theta$ is the weight assigned to stabilizing the output gap relative to inflation. x^* is the target level of the output gap, which stems from monopolistic competition and distortion in the steady state. Output subsidies are assumed to offset the monopolistic distortion so that the steady state is efficient, $x^* = 0$. As a result, in this analysis, there is no inflation bias but there is a stabilization bias due to suboptimal monetary policy and markup shocks, even if monetary policy is not constrained by the ZLB.

---

seller's desired markup is $\theta/(\theta-1)$. Moreover, $\alpha \in (0,1)$ is the share of firms keeping prices fixed each period, so the implied duration between price changes is $1/(1-\alpha)$.

2.2 Monetary Policy

I consider four monetary policy frameworks, namely a simple policy rule, two targeting rules, and optimal commitment policy. In each of these policy frameworks, the central bank faces a ZLB constraint and persistent errors in the measurement of economic conditions. Regardless of the ZLB constraint, the measurement errors lead to policy mistakes and therefore cause a deterioration in economic performance.

The first policy framework is an *inertial Taylor rule* subject to a ZLB constraint, along the lines of Taylor and Williams (2010),

$$i_t = \max\left[0, \phi_i i^u_{t-1} + (1-\phi_i)\left(r + \phi_\pi \pi^o_t + \phi_x x^o_t\right)\right], \tag{5}$$

where ϕ_π and ϕ_x are positive response coefficients on observed inflation, $\pi^o_t = \pi_t + e^\pi_t$, and the observed output gap, $x^o_t = x_t + e^x_t$, respectively. e^π_t and e^x_t represent noise shocks or measurement errors.[8] This rule incorporates smoothing in the behavior of the interest rate, through a positive value of the coefficient $\phi_i \in [0, 1)$. Moreover, i^u_{t-1} denotes an unconstrained or notional interest rate, the preferred setting of the policy rate in the previous period that would occur absent the ZLB constraint. Thus, the policy rate is kept below the notional interest rate following an episode when the ZLB is a binding constraint on policy.[9] This inertial Taylor rule is used for the calibration of the model (section 2.4).

The next two policy frameworks considered are targeting rules subject to a ZLB constraint. In other words, rather than following a simple policy rule, the central bank aims to stabilize a target variable by reoptimizing to the extent possible its policy decision ($i_t \geq 0$) in each period. One of the targeting rules considered is *strict inflation targeting*,

$$\pi^o_t = 0 \text{ subject to } i_t \geq 0, \tag{6}$$

[8]In the data, both inflation and output gaps are subject to persistent revisions (section 2.4). Thus, in the model, instead of using only one noise shock to reduce the number of state variables, both e^π_t and e^x_t are present for the policy rule to be consistent with the real-time data.

[9]Such an approach implies that the central bank compensates to some extent for the lost monetary stimulus due to the presence of the ZLB, even though the central bank does not commit to making up for past shortfalls from a nominal-level target.

where the central bank seeks to stabilize inflation without any concern for output stability and, therefore, transfers the burden of shocks onto output. This targeting rule does not involve any inertia in the setting of monetary policy, because the current policy decision disregards past economic conditions and past misses from the target.

The other targeting rule considered in this analysis is *nominal-GDP-level targeting*,

$$n_t^o = 0 \text{ subject to } i_t \geq 0, \tag{7}$$

where n_t^o is observed nominal GDP, $n_t^o = n_t + e_t^n$. Specifically, $n_t = p_t + y_t$ is actual nominal GDP measured as the log-deviation from a trend, and e_t^n is a noise shock. With this targeting rule, the central bank seeks to stabilize nominal GDP, as opposed to focusing entirely on inflation stability, which now requires the burden of shocks to be shared by inflation and output. This targeting rule involves inertia in the behavior of monetary policy because the current policy decision depends on the past price level, as $p_t \equiv p_{t-1} + \pi_t$.

Next, as a benchmark for the evaluation of these monetary policy rules, I use the *optimal commitment policy*. In such a policy framework, the central bank is assumed able and willing to fully commit to its policy announcements, to maximize the welfare of the representative household. In this ideal policy framework, the central bank's objective function is given by

$$\min_{i_t \geq 0} E_0 \sum_{t=0}^{\infty} \beta^t \left[(\pi_t^o)^2 + \lambda (x_t^o)^2 \right],$$

where the central bank seeks to stabilize to the extent possible inflation and the output gap, subject to a ZLB constraint. This objective function generally differs from the social welfare function, equation (4), because the central bank faces persistent errors in the measurement of inflation and the output gap. These measurement errors cause a deterioration in economic performance, regardless of the ZLB constraint.

In these four monetary policy frameworks, the exogenous noise shocks (e_t^π, e_t^x, e_t^n) are assumed to follow AR(1) stochastic processes, with first-order autocorrelation parameters $\rho_j \in [0, 1)$ for

$j = e^{\pi}, e^{x}, e^{n}$. Moreover, $\sigma_{\varepsilon j}\varepsilon_{jt}$ are the shock innovations buffeting the economy, which are independent across time and cross-sectionally, and normally distributed with mean zero and standard deviations $\sigma_{\varepsilon j} > 0$.

2.3 Equilibrium

At equilibrium, the policymaker chooses a policy based on a response function $\mathbf{y}(\mathbf{s}_t)$ and a state vector $\mathbf{s}_t$. The $\mathbf{s}_t$ includes the endogenous variables, the structural shocks, as well as the noise shocks affecting the central bank's observation of economic conditions. The corresponding expectations function is then

$$\mathbf{E}_t\mathbf{y}(\mathbf{s}_{t+1}) = \int \mathbf{y}(\mathbf{s}_{t+1}) f(\varepsilon_{t+1}) d(\varepsilon_{t+1}),$$

where $f(\cdot)$ is a probability density function of future innovations, both in the structural and noise shocks, which buffet the economy. In such a setting, an equilibrium is given by a response function and expectations function, $\mathbf{y}(\mathbf{s}_t)$ and $\mathbf{E}_t\mathbf{y}(\mathbf{s}_{t+1})$, which satisfy the equilibrium conditions, derived in section A.1 of the appendix.

Ignoring the existence of uncertainty about the future state of the economy, the model can be solved with standard numerical methods, as done in Orphanides and Wieland (2000), Reifschneider and Williams (2000), Williams (2009), Levin et al. (2010), Coibion, Gorodnichenko, and Wieland (2012), and Guerrieri and Iacoviello (2015), among others. When the ZLB threatens, however, the mere possibility of hitting the ZLB causes expectations of a future economic downturn and therefore prompts for adding policy stimulus today, as shown by Adam and Billi (2006, 2007), Nakov (2008), and others. In this analysis, as in Billi (2011a, 2017), I use a numerical procedure that accounts for the ZLB constraint and uncertainty about the evolution of the economy.[10]

2.4 Baseline Calibration

The model economy is calibrated to revised U.S. data for recent decades, as in Billi (2017), with monetary policy described by the

[10]See section A.2 of the appendix for a description of the algorithm used to solve the model.

inertial Taylor rule (5) that features prominently in Federal Reserve discussions. The values of the rule coefficients are taken from English, López-Salido, and Tetlow (2015), with ϕ_π set to 1.5, ϕ_x set to 1/4 (quarterly rates), and ϕ_i set to 0.85. The rule thus accounts for smoothing in the setting of the policy interest rate.

The values of the structural parameters are also standard in the related literature. Specifically, β is set to 0.993, to imply r equal to 3 percent annual. φ is set to 6.25.[11] The implied parameters κ and λ are then equal to 0.024 and 0.003 (quarterly), respectively. Regarding the structural shocks, $\rho_{y^n,u,v}$ are set to 0.8, to generate persistent effects on the economy. $\sigma_{y^n,v}$ are set to 0.8 percent (quarterly) to try to replicate respectively the volatility of output and nominal interest rates in the data, whereas σ_u is set to 0.05 percent (quarterly) to match the inflation volatility in the data.[12] Overall, as Billi (2017) showed, with the inertial Taylor rule and revised data, the model does a fairly good job in replicating the relevant features of recent U.S. data.[13]

[11] α is set to 0.66, so the duration between price changes $1/(1-\alpha)$ is three quarters. θ is set to 7.66, so the markup over marginal cost $\theta/(\theta-1)$ is 15 percent. Moreover, ω is set to 0.47.

[12] The inflation rate is measured as the continuously compounded rate of change in the seasonally adjusted personal consumption expenditures chain-type price index less food and energy (source: Bureau of Economic Analysis, BEA). Output is measured as the log-deviation from trend in seasonally adjusted gross domestic product (source: BEA). The output gap is calculated as the deviation of real gross domestic product from potential, as a fraction of potential using seasonally adjusted data (source: Congressional Budget Office). And the nominal interest rate is measured as the average effective federal funds rate (source: Federal Reserve Board). The sample period used to calibrate the structural shocks is the same as in Billi (2017), 1984:Q1–2014:Q4, which ensures the results are directly comparable. Moreover, extending the sample to the latest available data does not affect the good fit of the model to the data. Real-time and revised data are obtained from archival economic data available at the Federal Reserve Bank of St. Louis, https://alfred.stlouisfed.org.

[13] Still, output and inflation are somewhat less persistent in the model results than in the data because this basic model, for the sake of simplicity, does not allow for structural propagation mechanisms that give rise to output and inflation inertia. As a consequence, the stylized model understates the frequency and duration of ZLB episodes. With the inertial Taylor rule and revised data, the model predicts that the policy rate hits the ZLB about 4 percent of the time, and the expected duration of a ZLB episode is about four quarters (table 2). In actuality, the federal funds rate has been near the ZLB from the end of 2008 to

The calibration of the noise shocks is obtained fitting historical revisions of U.S. data, as done in Billi (2011b). Real-time estimates reflect information actually available to policymakers in each quarter, whereas revised estimates reflect information available at the end of the sample period. The difference between revised and real-time estimates is the historical revision of the data. $(\sigma_{\varepsilon e^{\pi}}, \sigma_{\varepsilon e^{x}}, \sigma_{\varepsilon e^{n}})$ are set to match the volatility of data revisions $(0.3, 1.7, 1.1)$ in percent annual. $(\rho_{e^{\pi}}, \rho_{e^{x}}, \rho_{e^{n}})$ are set to match the persistence of data revisions $(0.7, 0.85, 0.8)$.[14] Thus, reflecting historical revisions of the data in the calibrated model, the measurement errors are notably larger and more persistent for the output gap than for nominal GDP and inflation.

3. The Policy Evaluation

With the calibrated model, I study the extent to which persistent measurement errors and a ZLB constraint adversely affect the performance of the two targeting rules and inertial Taylor rule, relative to the optimal commitment policy. Exploring a range of calibrations for the supply and demand shocks buffeting the economy, I show that the ranking of these monetary policy rules depends crucially on the likelihood of hitting the ZLB constraint.

3.1 Response to Shocks

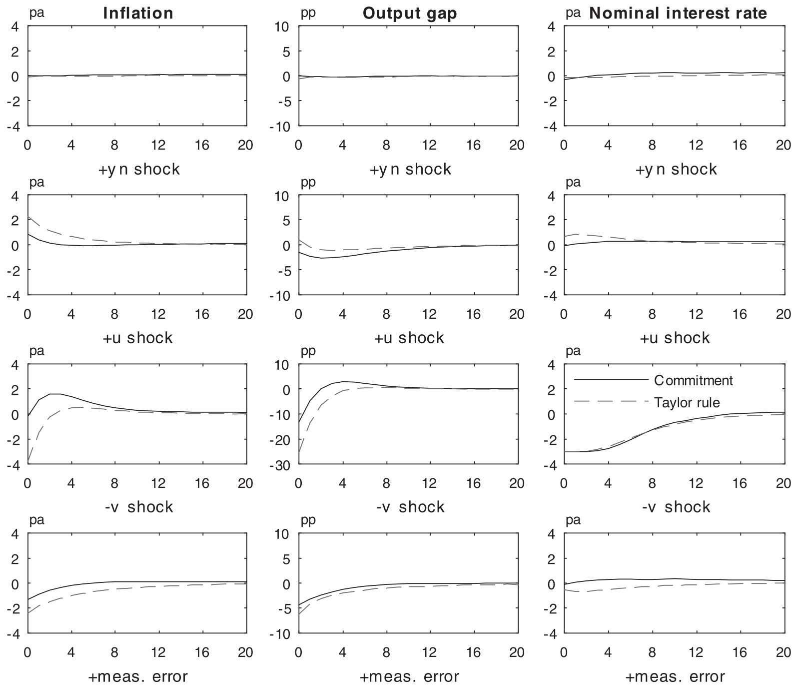
As the first step in the evaluation of the monetary policy rules, figures 1 and 2 show the evolution of the economy when hit by the different types of structural and noise shocks in the model.[15] The first figure displays responses under the optimal commitment

the end of 2015. See section 2.4 of Billi (2017) for further details of the model calibration and fit to the data.

[14]The half-lives of the noise shocks $log\,(0.5)\,/log\,(\rho_{e^{\pi}}, \rho_{e^{x}}, \rho_{e^{n}})$ are equal to $(1.9, 4.3, 3.1)$ quarters.

[15]Shown are expected paths after three-standard-deviation shocks, using the baseline calibration described in section 2.4. The expected paths are obtained by averaging across 10,000 stochastic simulations. Regarding the paths shown in figure 1, both inflation and the output gap are assumed to be mismeasured at the same time. If instead only inflation or the output gap is mismeasured, the resulting paths would be of similar shape but smaller size than the ones displayed in the figure.

Figure 1. Evolution of the Economy under Optimal Commitment or Taylor Rule

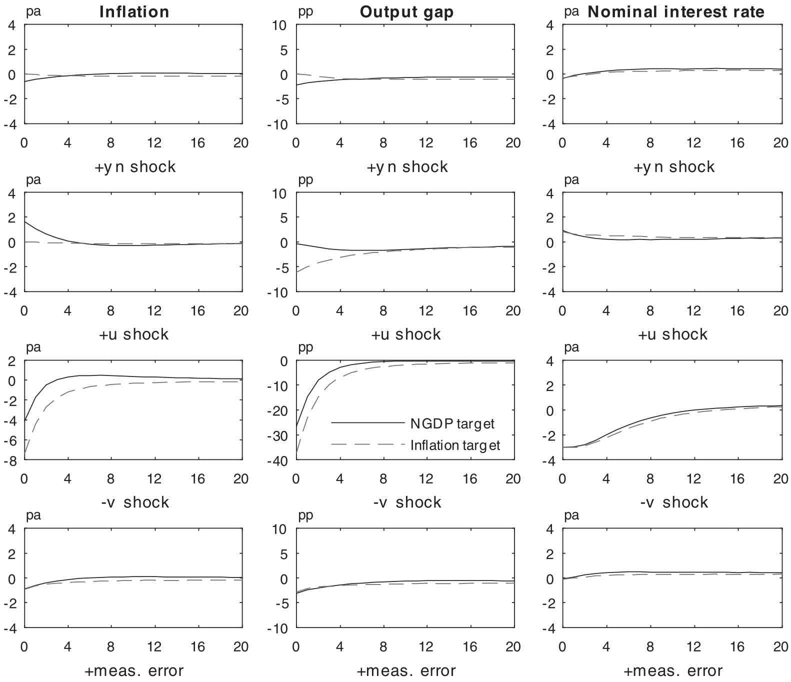


Notes: Shown are expected paths after three-standard-deviation shocks, using the baseline calibration of section 2.4. Values are expressed as percent annual (pa) or in percentage points (pp), in deviation from steady state.

policy and inertial Taylor rule, whereas the second figure displays responses under the two targeting rules. In both figures are shown the responses of inflation, the output gap, and the nominal interest rate.

Regarding the supply shocks, the top panel of the two figures shows the response to a positive technology shock, which implies slack in economic activity and downward pressure on prices. The outcome, however, depends on the monetary policy rule considered. With NGDP-level targeting (figure 2, solid line) both the output gap and inflation fall, whereas under the other three policy frameworks the economy is generally stabilized. In other words, in contrast to

Figure 2. Evolution under NGDP-Level Targeting or Strict Inflation Targeting



Notes: Shown are expected paths after three-standard-deviation shocks, using the baseline calibration of section 2.4. Values are expressed as percent annual (pa) or in percentage points (pp), in deviation from steady state.

the other policy frameworks, NGDP-level targeting fails to insulate the economy from technology shocks, conditional on no other shocks buffeting the economy.

The second panel shows the response to a positive markup shock, which implies upward pressure on prices and creates an inflation-output tradeoff for monetary policy. Facing such a tradeoff, under strict inflation targeting (figure 2, dashed line), inflation is completely stabilized and output falls notably, whereas under the other three policy frameworks inflation rises and output falls. The reason is that, in contrast to the other policy frameworks, strict inflation targeting transfers the burden of markup shocks onto output. The other policy frameworks require instead the burden of shocks to be shared by inflation and output.

Regarding the remaining shocks in the model, the third panel of the figures shows the response to a negative demand shock, which exerts downward pressure on output and prices. Given the size of the shock, under each policy framework, the weakness of the economy prompts the central bank to cut the nominal interest rate all the way to the ZLB. During the ZLB episode, both output and inflation fall to a greater extent under strict inflation targeting, compared with the other policy frameworks considered. The reason is that, as noted earlier, strict inflation targeting does not involve any inertia in the setting of monetary policy.

Finally, the bottom panel of the figures shows the response to a positive noise shock or measurement error, which implies the central bank incorrectly assumes there is upward pressure on prices. As a consequence of such a measurement error, under each policy framework considered, the central bank mistakenly tightens the stance of monetary policy and therefore causes inflation and output to fall.[16]

3.2 Economic Performance

The ability of the central bank to stabilize the economy is adversely affected by persistent measurement errors and a ZLB constraint. To illustrate, table 1 summarizes the performance of the optimal commitment policy, in four different cases.

In the first case there is neither measurement error nor ZLB in the model, in the second there is measurement error only, in the third instead there is the ZLB only, and in the fourth there are both measurement error and ZLB. The table reports for each case the expected frequency and duration of ZLB episodes, as well as the welfare loss due to business cycles.[17] As the table shows, taking into account either measurement error or the ZLB causes a deterioration in economic performance, as both inflation and output become more

[16] As the real interest rate (not shown) is higher, the stance of monetary policy is tightened.

[17] To calculate the welfare loss, first the value of the objective function (4) is obtained by averaging across 10,000 stochastic simulations, each 1,000 periods long after a burn-in period. This value is then converted into a permanent consumption loss, as explained in section A.3 of the appendix.

Table 1. Economic Performance under the Optimal Commitment Policy[a]

	ZLB Episodes		Welfare Loss[b]		
	Freq.[c]	Duration[d]	π	x	Tot.
No Measurement Error, No ZLB	0.0	0.0	0.01	0.02	0.03
Measurement Error Only	0.0	0.0	0.03	0.03	0.06
ZLB Only	14.5	2.9	0.03	0.05	0.08
Measurement Error and ZLB	11.3	3.1	0.05	0.06	0.11

[a]Baseline calibration of section 2.4.
[b]Permanent consumption loss in percentage points.
[c]Expected percent of the time at ZLB.
[d]Expected number of consecutive quarters at ZLB.

variable. However, economic performance is more severely affected by the ZLB.[18]

I now rank the monetary policy rules in terms of performance, relative to the optimal commitment policy, and also study whether the ranking depends on the measurement error and ZLB constraint. Table 2 summarizes the performance of each policy framework, again in the four cases. Each case is in a separate panel. In the first case, which is in the absence of measurement error and ZLB in the model, strict inflation targeting results in a smaller total welfare loss than NGDP-level targeting. With strict inflation targeting the burden of shocks is transferred onto output, but under NGDP-level targeting the shocks affect to a greater extent the volatility of inflation. Moreover, both these targeting rules perform better than the inertial Taylor rule. This ranking of the monetary policy rules is obtained also in the second case, which is in the presence of only measurement error in the model. However, because of the measurement error, strict inflation targeting is no longer able to transfer the entire burden of shocks onto output.

[18]The total welfare loss is 0.06 with measurement error only but rises to 0.08 with the ZLB only (table 1).

Table 2. Performance of the Monetary Policy Rules[a]

	ZLB Episodes		Welfare Loss[b]		
	Freq.[c]	**Duration[d]**	π	x	**Tot.**
No Measurement Error, No ZLB					
Commitment	0.0	0.0	0.01	0.02	0.03
NGDP Target	0.0	0.0	0.06	0.01	0.07
Inflation Target	0.0	0.0	0.00	0.05	0.05
Taylor Rule	0.0	0.0	0.21	0.25	0.46
Measurement Error Only					
Commitment	0.0	0.0	0.03	0.03	0.06
NGDP Target	0.0	0.0	0.07	0.02	0.09
Inflation Target	0.0	0.0	0.02	0.06	0.08
Taylor Rule	0.0	0.0	0.31	0.27	0.58
ZLB Only					
Commitment	14.5	2.9	0.03	0.05	0.08
NGDP Target	6.2	1.7	0.09	0.10	0.19
Inflation Target	11.9	2.3	0.13	0.25	0.38
Taylor Rule	4.5	3.4	0.23	0.29	0.52
Measurement Error and ZLB					
Commitment	11.3	3.1	0.05	0.06	0.11
NGDP Target	6.3	1.7	0.11	0.11	0.22
Inflation Target	11.6	2.3	0.15	0.26	0.41
Taylor Rule	3.8	3.4	0.33	0.32	0.65

[a]Baseline calibration of section 2.4.
[b]Permanent consumption loss in percentage points.
[c]Expected percent of the time at ZLB.
[d]Expected number of consecutive quarters at ZLB.

Turning to the third case, which is in the presence of the ZLB only, NGDP-level targeting now results in a smaller total welfare loss than strict inflation targeting. Moreover, due to the ZLB, strict inflation targeting is unable to transfer the entire burden of shocks onto output. The reason for this change in the ranking among the targeting rules is that, as noted earlier, strict inflation targeting does not involve any inertia in the setting of monetary policy, whereas under NGDP-level targeting the current policy decision depends on the past price level. Nevertheless, both targeting rules still perform better than the inertial Taylor rule. This ranking of the monetary policy rules remains the same when turning to the fourth case, which is in the presence of both measurement error and ZLB in the model.[19]

In summary, the implications of measurement error and a ZLB constraint for the ranking of the monetary policy rules are twofold. On the one hand, taking into account measurement error does not affect the ranking of the rules considered. On the other hand, taking into account the ZLB constraint inverts the ranking of the targeting rules, as NGDP-level targeting provides inertia in the setting of monetary policy and therefore outperforms strict inflation targeting when in the presence of the ZLB constraint.

3.3 Alternate Calibrations

The ranking of the monetary policy rules is affected by the likelihood of hitting the ZLB constraint. To illustrate, I modify the calibration of the supply and demand shocks, in the presence of both measurement error and ZLB in the model.[20]

I start by increasing the role of supply shocks, with table 3 summarizing the resulting performance of the policy frameworks. First, technology shocks are assumed to be substantially larger or more persistent than in the baseline calibration. However, as a comparison with the bottom panel of table 2 shows, this more prominent role assigned to technology shocks does not affect the volatility of the output gap and inflation and therefore leaves the policy ranking unchanged. In other words, after accounting for measurement error and ZLB, all four policy frameworks manage to insulate the economy from these bigger technology shocks. Second, markup shocks

[19]The result that with the baseline calibration of the model the two targeting rules perform better than the inertial Taylor rule should not be viewed as an argument against the use of simple policy rules, as this ranking depends on the calibration of the Taylor rule. For an illustration of this point, see section A.4 of the appendix.

[20]In tables 3 and 4, each type of structural shock is modified, either raising its standard deviation by 25 percent or increasing its persistence from 0.8 to 0.9, relative to the baseline calibration. Also shown in table 4, r is lowered from 3 to 2 percent annual, with β raised accordingly from 0.993 to 0.995.

Table 3. Performance of the Rules, Alternate Calibrations[a]

	ZLB Episodes		Welfare Loss[b]		
	Freq.[c]	Duration[d]	π	x	Tot.
Larger Technology Shocks ($\sigma_{y^n} = 1$)					
Commitment	11.1	3.1	0.05	0.06	0.11
NGDP Target	6.0	1.7	0.11	0.11	0.22
Inflation Target	11.4	2.3	0.15	0.26	0.41
Taylor Rule	3.7	3.3	0.33	0.32	0.65
More Persistent Technology Shocks ($\rho_{y^n} = 0.9$)					
Commitment	11.4	3.0	0.05	0.06	0.11
NGDP Target	5.9	1.7	0.11	0.11	0.22
Inflation Target	12.4	2.4	0.15	0.26	0.41
Taylor Rule	3.8	3.4	0.33	0.32	0.65
Larger Markup Shocks ($\sigma_u = 0.0625$)					
Commitment	11.3	3.1	0.06	0.07	0.13
NGDP Target	6.1	1.7	0.13	0.11	0.24
Inflation Target	11.2	2.3	0.16	0.29	0.45
Taylor Rule	3.7	3.3	0.39	0.33	0.72
More Persistent Markup Shocks ($\rho_u = 0.9$)					
Commitment	11.2	3.1	0.05	0.07	0.12
NGDP Target	5.6	1.7	0.12	0.12	0.24
Inflation Target	12.8	2.4	0.15	0.26	0.41
Taylor Rule	4.6	3.6	0.40	0.34	0.74

[a]With measurement error and ZLB in the model.
[b]Permanent consumption loss in percentage points.
[c]Expected percent of the time at ZLB.
[d]Expected number of consecutive quarters at ZLB.

are assumed to be substantially larger or more persistent than in the baseline calibration. This more prominent role given to markup shocks generally leads to higher volatility of the output gap and inflation, but the policy ranking is still unchanged.

Table 4. Alternate Calibrations[a]

	ZLB Episodes		Welfare Loss[b]		
	Freq.[c]	**Duration**[d]	π	x	**Tot.**
Larger Demand Shocks ($\sigma_v = 1$)					
Commitment	17.9	3.8	0.10	0.10	0.20
NGDP Target	10.0	2.0	0.16	0.25	0.41
Inflation Target	16.5	2.7	0.52	0.71	1.23
Taylor Rule	7.6	3.9	0.42	0.55	0.97
More Persistent Demand Shocks ($\rho_v = 0.9$)					
Commitment	13.2	5.4	0.10	0.06	0.16
NGDP Target	6.4	2.4	0.18	0.21	0.39
Inflation Target	12.1	3.5	5.03	3.60	8.63
Taylor Rule	5.7	3.7	0.37	0.42	0.79
Lower Steady-State Interest Rate ($r = 0.5$)					
Commitment	19.7	4.1	0.10	0.09	0.19
NGDP Target	10.6	2.0	0.16	0.24	0.40
Inflation Target	19.2	2.9	0.66	0.80	1.46
Taylor Rule	10.4	4.3	0.38	0.42	0.80

[a]With measurement error and ZLB in the model.
[b]Permanent consumption loss in percentage points.
[c]Expected percent of the time at ZLB.
[d]Expected number of consecutive quarters at ZLB.

Next, I assign instead a more prominent role to demand shocks, with table 4 summarizing the resulting performance of the policy frameworks. First, demand shocks are assumed to be substantially larger or more persistent than in the baseline calibration. As a comparison with the bottom panel of table 2 shows, under each policy framework, inflation and output generally become more variable. But this deterioration in economic performance is substantially larger under strict inflation targeting, compared with the other policy frameworks. The reason is that, as noted earlier, strict inflation targeting does not involve any inertia in the setting of monetary policy. As a result, strict inflation targeting now results in a larger total welfare loss than the other policy rules. Finally, the steady-state interest rate is substantially lower than in the baseline calibration, which implies that monetary policy is more severely constrained by the ZLB. Again strict inflation targeting is outperformed by the other policy rules.

In summary, these changes to the calibration imply the following for the ranking of the monetary policy rules. Even if the economy is hit by supply shocks that are substantially larger or more persistent than in the baseline calibration, the policy ranking is not affected. But if the role of demand shocks is more prominent than in the baseline calibration, and therefore monetary policy is more severely constrained by the ZLB, then strict inflation targeting is outperformed by the other policy rules considered in this analysis.

4. Concluding Remarks

Policymakers often use the output gap to guide monetary policy, even though inflation and nominal GDP are measured more accurately in real time than the output gap. Employing a small New Keynesian model, which offers a clear illustration of the tradeoffs faced by a central bank during ZLB episodes, this article compares the performance of monetary policy rules that are robust to persistent errors in the measurement of economic conditions.

The analysis shows that, in the absence of the ZLB, the central bank should focus on stabilizing inflation rather than nominal GDP. But when the ZLB is present, a policy that seeks to stabilize the level of nominal GDP improves substantially the tradeoffs faced by the central bank. Still, the analysis is conducted in a stylized model that does not include an explicit role for balance sheet policies, nor monetary policies based on monetary aggregates that have the potential to circumvent the ZLB constraint. See Belongia and Ireland (2017) for a discussion. It would be interesting to extend the analysis to include such features in future research.

Appendix

This appendix is organized in four sections. Section A.1 derives the equilibrium conditions of the model. Section A.2 describes the numerical procedure used to solve the model. Section A.3 explains the calculation of the permanent consumption loss. Section A.4 provides additional results about the evaluation of the inertial Taylor rule relative to the baseline calibration of the model.

A.1 Equilibrium Conditions

I first derive the equilibrium conditions and then summarize them in a table.

Optimal Commitment Policy. The problem can be written as

$$V(\mathbf{s}_t) = \min \left[(\pi_t + e_t^{\pi})^2 + \lambda (y_t - y_t^n + e_t^x)^2 + \beta E_t V(\mathbf{s}_{t+1}) \right]$$

$$\text{subject to (1), (3) and } i_t \geq 0.$$

Write the period Lagrangian:

$$\begin{aligned} L_t = {} & (\pi_t + e_t^{\pi})^2 + \lambda (y_t - y_t^n + e_t^x)^2 + \beta E_t V(\mathbf{s}_{t+1}) \\ & + m_{1t} [\pi_t - \kappa (y_t - y_t^n) - u_t] - m_{1t-1} \pi_t \\ & + m_{2t} [y_t + \varphi (i_t - r - v_t)] - \beta^{-1} m_{2t-1} (y_t + \varphi \pi_t). \end{aligned}$$

The Kuhn-Tucker conditions are

$$0 = \partial L_t / \partial \pi_t = 2 (\pi_t + e_t^{\pi}) + m_{1t} - m_{1t-1} - \beta^{-1} \varphi m_{2t-1} \quad \text{(A.1)}$$

$$0 = \partial L_t / \partial y_t = 2\lambda (y_t - y_t^n + e_t^x) - \kappa m_{1t} + m_{2t} - \beta^{-1} m_{2t-1} \quad \text{(A.2)}$$

$$0 = \partial L_t / \partial i_t \cdot i_t = \varphi m_{2t} \cdot i_t, \quad m_{2t} \geq 0, \quad i_t \geq 0. \quad \text{(A.3)}$$

The equilibrium conditions are summarized in table A.1.

A.2 Numerical Procedure

I find a numerical solution, as in Billi (2011a, 2017), as a fixed point in the equilibrium conditions. To do so, the state vector is discretized into a grid of interpolation nodes, with a support of ± 4 standard deviations for each state variable, which is large enough to avoid erroneous extrapolation. If the state is not on this grid, the response function is evaluated with multilinear interpolation. The approximation residuals are evaluated at a finer grid, to ensure the accuracy of the results. The expectations function is evaluated

Table A.1. Summary of Equilibrium Conditions

Policy	Equilibrium Conditions	State Vector $\mathbf{s}_t$
Optimal Commitment	(1), (3), and (A.1)–(A.3)	$(y_t^n, u_t, v_t, e_t^\pi, e_t^x, m_{1t-1}, m_{2t-1})$
Inertial Taylor Rule	(1), (3), and (5)	$(y_t^n, u_t, v_t, e_t^\pi, e_t^x, i_{t-1}^u)$
Strict Inflation Target	(1), (3), and (6)	$(y_t^n, u_t, v_t, e_t^\pi)$
NGDP-Level Target	(1)–(3) and (7)	$(y_t^n, u_t, v_t, e_t^n, p_{t-1})$

with Gaussian-Hermite quadrature. The initial guess is the linearized solution that ignores the ZLB constraint. This numerical procedure is coded in MATLAB. Replication files are available from the author upon request.

A.3 Permanent Consumption Loss

I obtain the permanent consumption loss as in Billi (2011a, 2017). The expected lifetime utility of the representative household is validly approximated by

$$E_0 \sum_{t=0}^{\infty} \beta^t U_t = \frac{U_c \overline{C}}{2} \frac{\alpha\theta\,(1+\omega\theta)}{(1-\alpha)\,(1-\alpha\beta)} L, \tag{A.4}$$

where $\overline{C}$ is steady-state consumption; $U_c > 0$ is steady-state marginal utility of consumption; and $L \geq 0$ is the value of objective function (4).

At the same time, a steady-state consumption loss of $\mu \geq 0$ causes a utility loss of

$$E_0 \sum_{t=0}^{\infty} \beta^t U_c \overline{C} \mu = \frac{1}{1-\beta} U_c \overline{C} \mu. \tag{A.5}$$

Equating the right sides of (A.4) and (A.5) gives

$$\mu = \frac{1-\beta}{2} \frac{\alpha\theta\,(1+\omega\theta)}{(1-\alpha)\,(1-\alpha\beta)} L.$$

Table A.2. Performance of Inertial Taylor Rule

		ZLB Episodes		Welfare Loss[a]		
ϕ_x	ϕ_π	Freq.[b]	Duration[c]	π	x	Tot.
0.250	1.5	3.8	3.4	0.33	0.32	0.65
0.125	1.5	2.6	3.1	0.40	0.53	0.93
0.250	5.0	5.7	3.6	0.14	0.24	0.38

[a]Permanent consumption loss in percentage points.
[b]Expected percent of the time at ZLB.
[c]Expected number of consecutive quarters at ZLB.

A.4 Evaluation of the Taylor Rule

With measurement error and ZLB in the model, table A.2 shows the performance of the inertial Taylor rule (5) for alternate values of its response coefficients compared with the baseline calibration of the model. The first line is the model outcome using the baseline calibration (that is, the same outcome as in the bottom panel of table 2).

This table illustrates two results about the response coefficients. First, a weaker response to the observed output gap than in the baseline causes a deterioration in economic performance, as inflation and output become more variable. The reason is that the weaker response to the output gap worsens the inflation-output tradeoff faced by the central bank. If ϕ_x is lowered from 0.25 to 0.125, the total welfare loss under the inertial Taylor rule increases from 0.65 to 0.93. Thus, even though the output gap is subject to large and persistent revisions, it is desirable for monetary policy to respond to a certain extent to the output gap.

Second, as the table also shows, a stronger response to observed inflation results in an improvement in economic performance. The inertial Taylor rule can even outperform strict inflation targeting but still performs worse than NGDP-level targeting. If ϕ_π is raised from 1.5 to 5, the total welfare loss under the inertial Taylor rule falls from 0.65 to 0.38. Instead under the targeting rules, the total welfare loss is 0.41 with strict inflation targeting but only 0.22 with NGDP-level targeting (bottom panel of table 2). Values of ϕ_π above 5 are not considered because they are typically viewed as impractically high.

References

Adam, K., and R. M. Billi. 2006. "Optimal Monetary Policy under Commitment with a Zero Bound on Nominal Interest Rates." *Journal of Money, Credit and Banking* 38 (7): 1877–1905.

———. 2007. "Discretionary Monetary Policy and the Zero Lower Bound on Nominal Interest Rates." *Journal of Monetary Economics* 54 (3): 728–52.

Aoki, K. 2003. "On the Optimal Monetary Policy Response to Noisy Indicators." *Journal of Monetary Economics* 50 (3): 501–23.

———. 2006. "Optimal Commitment Policy under Noisy Information." *Journal of Economic Dynamics and Control* 30 (1): 81–109.

Bean, C. R. 1983. "Targeting Nominal Income: An Appraisal." *Economic Journal* 93 (372): 806–19.

Belongia, M. T., and P. N. Ireland. 2017. "Circumventing the Zero Lower Bound with Monetary Policy Rules Based on Money." *Journal of Macroeconomics* 54 (Part A): 42–58.

Billi, R. M. 2011a. "Optimal Inflation for the U.S. Economy." *American Economic Journal: Macroeconomics* 3 (3): 29–52.

———. 2011b. "Output Gaps and Monetary Policy at Low Interest Rates." *Economic Review* (Federal Reserve Bank of Kansas City) 96 (First Quarter): 5–29.

———. 2017. "A Note on Nominal GDP Targeting and the Zero Lower Bound." *Macroeconomic Dynamics* 21 (8): 2138–57.

Boehm, C. E., and C. L. House. 2014. "Optimal Taylor Rules in New Keynesian Models." NBER Working Paper No. 20237.

Clark, T. E. 1994. "Nominal GDP Targeting Rules: Can They Stabilize the Economy?" *Economic Review* (Federal Reserve Bank of Kansas City) 79 (Third Quarter): 11–25.

Coibion, O., Y. Gorodnichenko, and J. Wieland. 2012. "The Optimal Inflation Rate in New Keynesian Models: Should Central Banks Raise Their Inflation Targets in Light of the Zero Lower Bound?" *Review of Economic Studies* 79 (4): 1371–1406.

Ehrmann, M., and F. Smets. 2003. "Uncertain Potential Output: Implications for Monetary Policy." *Journal of Economic Dynamics and Control* 27 (9): 1611–38.

English, W. B., J. D. López-Salido, and R. J. Tetlow. 2015. "The Federal Reserve's Framework for Monetary Policy: Recent

Changes and New Questions." *IMF Economic Review* 63 (1): 22–70.

Frankel, J. 2013. "Nominal-GDP Targets, without Losing the Inflation Anchor." In *Is Inflation Targeting Dead? Central Banking After the Crisis,* ed. L. Reichlin, and R. Baldwin, 90–94. London: CEPR, VoxEU.org.

Garín, J., R. Lester, and E. Sims. 2016. "On the Desirability of Nominal GDP Targeting." *Journal of Economic Dynamics and Control* 69 (August): 21–44.

Guerrieri, L., and M. Iacoviello. 2015. "OccBin: A Toolkit for Solving Dynamic Models with Occasionally Binding Constraints Easily." *Journal of Monetary Economics* 70 (March): 22–38.

Gust, C. J., B. K. Johannsen, and J. D. López-Salido. 2017. "Monetary Policy, Incomplete Information, and the Zero Lower Bound." *IMF Economic Review* 65 (1): 37–70.

Hall, R. E., and N. G. Mankiw. 1994. "Nominal Income Targeting." In *Monetary Policy*, ed. N. G. Mankiw, 71–93 (chapter 2). Chicago and London: University of Chicago Press.

Hatzius, J., and S. J. Stehn. 2011. "The Case for a Nominal GDP Level Target." *US Economics Analyst* (Goldman Sachs) 11/41, October 14.

———. 2013. "A Nominal GDP Level Target in All but Name." *US Economics Analyst* (Goldman Sachs) 13/03, January 20.

Jensen, H. 2002. "Targeting Nominal Income Growth or Inflation?" *American Economic Review* 92 (4): 928–56.

Levin, A., D. López-Salido, E. Nelson, and T. Yun. 2010. "Limitations on the Effectiveness of Forward Guidance at the Zero Lower Bound." *International Journal of Central Banking* 6 (1, March): 143–89.

McCallum, B. T. 1988. "Robustness Properties of a Rule for Monetary Policy." *Carnegie-Rochester Conference Series on Public Policy* 29 (1): 173–204.

Meade, J. E. 1978. "The Meaning of 'Internal Balance'." *Economic Journal* 88 (351): 423–35.

Nakov, A. 2008. "Optimal and Simple Monetary Policy Rules with Zero Floor on the Nominal Interest Rate." *International Journal of Central Banking* 4 (2, June): 73–127.

Orphanides, A. 2001. "Monetary Policy Rules Based on Real-Time Data." *American Economic Review* 91 (4): 964–85.

———. 2003. “Monetary Policy Evaluation with Noisy Information.” *Journal of Monetary Economics* 50 (3): 605–31.

Orphanides, A., R. D. Porter, D. Reifschneider, R. Tetlow, and F. Finan. 2000. “Errors in the Measurement of the Output Gap and the Design of Monetary Policy.” *Journal of Economics and Business* 52 (1–2): 117–41.

Orphanides, A., and S. van Norden. 2002. “The Unreliability of Output-Gap Estimates in Real Time.” *Review of Economics and Statistics* 84 (4): 596–83.

Orphanides, A., and V. Wieland. 2000. “Efficient Monetary Policy Design Near Price Stability.” *Journal of the Japanese and International Economies* 14 (4): 327–65.

Reifschneider, D., and J. C. Williams. 2000. “Three Lessons for Monetary Policy in a Low-Inflation Era.” *Journal of Money, Credit and Banking* 32 (4): 936–66.

Rudebusch, G. D. 2002. “Assessing Nominal Income Rules for Monetary Policy with Model and Data Uncertainty.” *Economic Journal* 112 (479): 402–32.

Smets, F. 2002. “Output Gap Uncertainty: Does It Matter for the Taylor Rule?” *Empirical Economics* 27 (1): 113–29.

Sumner, S. B. 2011. “Re-Targeting the Fed.” *National Affairs* 41 (Fall): 79–96.

———. 2014. “Nominal GDP Targeting: A Simple Rule to Improve Fed Performance.” *Cato Journal* 34 (2): 315–37.

Svensson, L. E. O. 2010. “Monetary Policy and Financial Markets at the Effective Lower Bound.” *Journal of Money, Credit and Banking* 42 (s1): 229–42.

Svensson, L. E. O., and M. Woodford. 2003. “Indicator Variables for Optimal Policy.” *Journal of Monetary Economics* 50 (3): 691–720.

———. 2004. “Indicator Variables for Optimal Policy under Asymmetric Information.” *Journal of Economic Dynamics and Control* 28 (4): 661–90.

Taylor, J. B. 1985. “What Would Nominal GNP Targeting Do to the Business Cycle?” *Carnegie-Rochester Conference Series on Public Policy* 22 (1): 61–84.

Taylor, J. B., and J. C. Williams. 2010. “Simple and Robust Rules for Monetary Policy.” In *Handbook of Monetary Economics*, Vol.

3, ed. B. M. Friedman, and M. Woodford, 829–59 (chapter 15). Amsterdam: Elsevier B.V.

Tobin, J. 1980. "Stabilization Policy Ten Years After." *Brookings Papers on Economic Activity* 11 (1): 19–90.

Walsh, C. E. 2003. "Speed Limit Policies: The Output Gap and Optimal Monetary Policy." *American Economic Review* 93 (1): 265–78.

West, K. D. 1986. "Targeting Nominal Income: A Note." *Economic Journal* 96 (384): 1077–83.

Williams, J. C. 2009. "Heeding Daedalus: Optimal Inflation and the Zero Lower Bound." *Brookings Papers on Economic Activity* 40 (2): 1–45.

Woodford, M. 2010. "Optimal Monetary Stabilization Policy." In *Handbook of Monetary Economics*, Vol. 3, ed. B. M. Friedman and M. Woodford, 723–828 (chapter 14). Amsterdam: Elsevier B.V.

———. 2012. "Methods of Policy Accommodation at the Interest-Rate Lower Bound." In *The Changing Policy Landscape*. Proceedings of the 2012 Economic Policy Symposium sponsored by the Federal Reserve Bank of Kansas City, held in Jackson Hole, Wyoming, August 30–September 1.

———. 2013. "Inflation Targeting: Fix It, Don't Scrap It." In *Is inflation Targeting Dead? Central Banking After the Crisis,* ed. L. Reichlin and R. Baldwin, 74–89. London: CEPR, VoxEU.org.

Does Central Bank Transparency and Communication Affect Financial and Macroeconomic Forecasts?*

Thomas Lustenberger[a] and Enzo Rossi[b]
[a]Swiss Financial Market Supervisory Authority (FINMA) and University of Basel
[b]Swiss National Bank and University of Zurich

In a large sample of countries across different geographic regions and over a long period of time, we find limited country- and variable-specific effects of central bank transparency on forecast accuracy and their dispersion among a large set of professional forecasts of financial and macroeconomic variables. More communication even increases forecast errors and dispersion.

JEL Codes: C23, C53, E37, E58, D8.

1. Introduction

Until not so long ago, central bankers believed that monetary policy decisions should take the markets by surprise in order to achieve maximum impact. In the last two decades, there has been a shift to

*We would like to thank the co-editor of the journal, Boragan Aruoba. We also thank Nicolas Albacete, Alex Cukierman, Beat Hintermann, Bo Honoré, Massimiliano Marcellino, Cyril Monnet, Marcel Savioz, Heinz Zimmermann, and three anonymous referees for valuable comments and suggestions. In addition, we profited from the comments of participants of the SNB Brown Bag Seminar, the Economics Lunch of the University of Basel, the BuBa-OeNB-SNB Workshop 2017, the SNB-BIS Workshop 2018, and the EABCN Conference 2018. We also thank Bettina Eberhard (BIS) and Paul Moser-Boehm (BIS) as well as Simon Dépraz for their help in collecting the central bank communication measure. A prior version of this essay was published as SNB Working Paper 2017-12 in October 2017 and WWZ Discussion Paper 2018-06 in January 2018. This research project was conducted while the first author was affiliated with the SNB and the University of Basel. The views expressed in this paper are those of the authors and do not necessarily reflect those of the SNB or FINMA. Author e-mails: thomas.lustenberger@me.com and enzo.rossi@snb.ch.

a policy of increasing transparency with respect to goals, strategies, and the basis on which decisions are made. This development is closely linked to increased central bank independence (CBI), which calls for a counterbalance in the form of transparency and accountability. Economic benefits are deemed another key reason for enhanced transparency. As summarized by Freedman and Laxton (2009), it is generally believed in the central banking community that providing more information about monetary policy may increase its effectiveness. This view is based on theoretical and empirical research that emphasizes the importance of expectations about monetary policy as a key element in determining interest rates and other asset prices.[1] By bringing market behavior in line with monetary policy objectives, the likelihood of sharply differing views on policy actions is reduced. In turn, more certainty about when the central bank will set the policy rate and its magnitude can reduce the volatility of market interest rates, increase the central bank's leverage over longer-term interest rates, and smoothen the incorporation of policy actions into asset prices. Similarly, Blinder (2007) argues that the major purpose of communicating with the markets is to condition expectations about future monetary policy.

To transmit the views of the central bank to the public and to markets, an improvement in the effectiveness of monetary policy through greater transparency requires proactive and well-planned communication. Hence, a great deal of attention has been paid to the way central banks present their key messages (for instance, Blinder et al. 2008 and Haldane 2017). It is expected that central banks will communicate more actively after than before the crisis (Blinder et al. 2017).

Parallel to the shifts in the practice of central banking toward more frequent communication and greater transparency, an expanding body of literature has emerged. In theory, both positive and negative effects are likely. Empirical work has been focused on inflation, financial markets, or private forecasts. A general finding is that transparency and communication reduce volatility in financial markets, enhance the predictability of upcoming rate decisions, and help achieve the monetary policy goals, vindicating the switch to greater openness in central banking.

[1] Survey forecasts are an approximation for expectations in an economy.

Our paper contributes specifically to the empirical relation between transparency, communication, central bank design, and private forecasts. In a broader context, we provide empirical results for the theoretical literature on the social value of information (see, for instance, Morris and Shin 2002, Hellwig and Veldkamp 2009, and Lorenzoni 2010).

We run panel regressions to examine whether increased transparency and intensified communication by central banks affect the quality and the cross-sectional distribution of forecasts. The question posed is closest to those of Middeldorp (2011), Dovern, Fritsche, and Slacalek (2012), Ehrmann, Eijffinger, and Fratzscher (2012), Neuenkirch (2013), and Naszodi et al. (2016). However, we extend the analysis along various important dimensions and provide compelling evidence that is in contrast with the literature.

First, while in previous work transparency and communication have not been defined specifically, our data set allows us to make a clear distinction between the two dimensions.[2] Transparency covers broadly the strategic, long-term orientation of the central bank toward openness which is subject to only a few changes over time, such as policy objectives, the publication of forecasts, voting records, etc. Communication, on the other hand, focuses on the tactical, flexible aspects in the provision of more detailed and fine-tuned information to the public and the markets.

[2]For instance, Faust and Svensson (2001) equate central bank transparency with the ease with which the public can deduce central bank goals and intentions from observable data. In Amato, Morris, and Shin (2002) greater transparency is achieved by expanding the modes of communication and the amount of information revealed to the public. Winkler (2002) emphasizes that transparency needs to be better defined before it can be debated. More information divulged by the central bank does not necessarily imply a greater understanding on the part of the public, part of his definition of transparency. In Chortareas, Stasavage, and Sterne (2002) transparency relates to the detail in which central banks publish economic forecasts. Blinder et al. (2008) define central bank communication more broadly as the provision of information to the public regarding the objectives of monetary policy, the monetary policy strategy, the economic outlook, and the outlook for future policy decisions. Ehrmann, Eijffinger, and Fratzscher (2012) draw a distinction between central bank transparency and communication. Transparency is measured by the announcement of a quantified inflation objective, the overall index provided by Dincer and Eichengreen (2010) as well as the subindex related to economic transparency. The last measure for transparency is the publication of internal forecasts, which they label central bank communication.

Second, we are the first to document an increase in the number of speeches held by central banks over time using the Bank for International Settlements (BIS) central bankers' speeches database. Based on it we compiled a new measure of central bank communication. It measures communication directly and comprehensively by the number of speeches. The number of speeches allows us to question the benefits of intensive central bank communication.

Third, we highlight an econometric issue in the estimation method that does not seem to have been recognized in the literature, and we offer an accurate alternative. The issue is that residuals follow a pattern when the dependent variable (either absolute cross-sectional mean forecast error or cross-sectional standard deviation) is not taken in log. There is a clear-cut lower bound to the value residuals assumed in the setting without log.

Fourth, we create an exceptionally large panel of 73 countries from all world regions and observations from 1998 to 2014 for financial and macroeconomic data. The question is whether there is a one-size-fits-all policy for central bank communication and transparency. Unlike previous studies, which have in common a limited number of advanced economies (at most around 30) and relatively short periods of observations and forecast variables, the data set we compiled allows us to widen the scope of the inquiry in terms of the number of countries, their heterogeneity, the period of investigation, and the variables to be forecasted.

Finally, our data set also allows us to account for several important economic events, such as the Great Moderation, the financial crisis, and the global recession, as well as the substantial modifications to central bank practices in their wake—in particular, forward guidance. We also analyze the effect of inflation targeting.

Overall, the evidence presented in this paper suggests a more balanced conclusion about the merits of communication and transparency in enhancing the predictability of monetary policy than has been reported in the literature.

First, the evidence for communication is uniform and quite compelling: more-frequent communication increases both forecast errors and their dispersion. The increased central bank communication seems to have resulted in cacophony and did not help investors and academics improve their macroeconomic forecasts. We link this result with the discussion about optimal monetary policy committees' size and form.

Second, while more-frequent communication increases both forecast errors and their dispersion, we find hardly any evidence, in contrast to previous papers, that transparency, defined as the strategic component of a central bank's openness, improves the accuracy of private forecasts. At best, the impact is ambiguous. However, if it is significant, transparency tends to reduce the forecast heterogeneity especially of inflation and interest rates.

However, it is important to distinguish the pre-crisis from the post-crisis period. In fact, before the financial crisis, more intense communication lowered the errors in inflation and short-term interest rate forecasts. After the crisis, we find strong evidence that more communication increased the errors in forecasts of inflation and short-term interest rate forecasts and also increased the dispersion of real GDP and long-term interest rate forecasts. Increased transparency had a beneficial effect on the accuracy and dispersion of forecasts before the crisis, but this effect notably weakened after the crisis.

Third, additional analysis provides information about other factors that have an effect on the precision and distribution of forecasts. (i) Central bank instability is associated with less-accurate current-year forecasts of inflation and short-term interest rates 12 months ahead. (ii) Market uncertainty increases strongly forecast dispersion across all financial and macroeconomic variables. (iii) The zero-lower-bound constraint tends to reduce forecast errors of short-term interest rates, but has no effect on long-term interest rate forecasts. (iv) An inflation-targeting regime has some interesting implications for forecast precision and homogeneity. While more communication by inflation targeters improves the accuracy of short-term interest rate forecasts, greater transparency of inflation targeters not only worsens forecasts' accuracy but also tends to offset an otherwise positive effect on the heterogeneity of inflation and interest rate forecasts. (v) The next set of results is related to explicit forward guidance, as adopted by some central banks in the follow-up to the financial crisis. The results show particularly that more speeches gave rise to less-accurate long-term interest rate forecasts.

Fourth, we add to the robustness of the evidence by confirming it across a variety of additional analyses.

A caveat is in order. Whether more or less communication or whether the degree of transparency should be increased or lowered cannot be definitely answered in our framework. Our paper only studies the effect of communication and transparency on forecast accuracy and dispersion. Although the effect of communication and transparency on this dimension is important, there may be many other beneficial (or harmful) effects of giving public speeches or being transparent on, for instance, accountability, the public's understanding of monetary policy, and trust in the central bank.

The rest of the paper is structured as follows. In section 2, we review the related literature. Section 3 describes the data underlying the empirical analysis. In section 4, we explain our estimation strategy and present the results. Section 5 offers a series of robustness checks. Section 6 discusses possible policy implications. Section 7 concludes.

2. Literature

There exists a large body of literature on the effects of central bank transparency, communication, and optimal central bank design.[3] The empirical evidence suggests overall beneficial effects. One branch of literature analyzes the effect of transparency on the predictability of monetary policy in the financial markets. A number of studies suggest that increased monetary policy transparency may have contributed to an increased ability of financial markets to forecast future monetary policy actions. Most of this research has used information from the Treasury bill markets as well as the markets for federal funds and Eurodollar futures, and it focuses on a relatively short-run horizon, from one day out to six months.

Three approaches have been pursued. One approach investigates the reaction of market prices to central bank decisions. Little reaction means the decision has been priced in correctly, suggesting high predictability. Evidence has been reported on this topic (see, for instance, Ranaldo and Rossi 2010 and Wilhelmsen and Zaghini

[3]See Van Der Cruijsen, Eijffinger, and Hoogduin (2010) for an overview of the transparency literature, Blinder et al. (2008) for a survey on communication, Geraats (2006) for an overview of the practice of monetary policy transparency, and Blinder (2004) for central bank design.

2011). The second approach is based on the accuracy of expectations priced into the yield curve or futures. Here, too, findings suggest that transparency leads to improved predictability (for instance, Kuttner 2001 and Lange, Sack, and Whitesell 2003). The third approach examines forecasts and/or the determinants of disagreement among forecasters. Swanson (2006) finds that with the increased transparency of the Federal Reserve, the private-sector forecasts of interest rates have become more precise, both by improving the average quality of forecasts and by reducing their dispersion across forecasters. In line with this, Sellon (2008) finds that more-explicit guidance on interest rates led to an improvement in private-sector forecasts.

The evidence stretches beyond the United States. Middeldorp (2011) analyzes the connection between the transparency and predictability of short-term interest rates for 24 countries between 1998 and 2005. Higher transparency lowers the errors private agents make in forecasting short-term interest at the three-month horizon, and it lowers the standard deviation. Dovern, Fritsche, and Slacalek (2012) investigate determinants of disagreement in expectations of seven key economic indicators in the G-7 countries from 1989 to 2006. In line with the literature (Mankiw, Reis, and Wolfers 2003), the measure of cross-sectional dispersion is the interquartile range of forecasts in a given country and month. While disagreement about economic activity intensifies strongly during recessions, disagreement about prices is considerably lower under independence of the central bank.[4] Based also on the interquartile range of forecasts, Ehrmann, Eijffinger, and Fratzscher (2012) examine whether transparency and communication have led to more-aligned views in the forecasts of macroeconomic variables in 12 advanced economies from 1990 to 2008. While transparency and communication reduce dispersion among professional forecasts, there is some evidence of diminishing marginal effects of increases in (economic) transparency. Naszodi et al. (2016) expand the analysis of Ehrmann, Eijffinger, and Fratzscher (2012) by enlarging the panel to 26 countries and by assessing both the degree of forecasting disagreement and its accuracy. Their results suggest that transparency results in better forecasts by mitigating uncertainty.

[4] Mankiw, Reis, and Wolfers (2003) show that a sticky-information model can generate a degree of disagreement among agents.

To sum up, the empirical literature provides support for the view that transparency is beneficial in the sense that survey forecasts are more aligned with each other and forecast errors decline. The evidence corroborates the general view that enhancing transparency improves the predictability of central banks.

Theoretical papers reach more a nuanced conclusion. On the one hand, more openness may reduce uncertainty about central banks' intentions and their future actions. On the other hand, by attempting to be as open as possible, they may give the impression that they know more than they do. This is a critical issue if transparency and communication serve as a coordination device among economic agents, thereby generating the possibility that agents rely too much on the utterances of central banks. This is what Morris and Shin (2002) argue can happen. Svensson (2006) disagrees with some of their conclusions. Subsequent research could not settle the matter.[5]

Restricting transparency could be worth considering for other reasons. For instance, the seminal paper by Cukierman and Meltzer (1986) argued that ambiguity enables monetary authorities to generate surprise inflation and stimulate economic activity. King (2000) notes that a central bank should be highly transparent about its monetary policy reaction function and its target. Beyond that, it should avoid creating news itself. Too much transparency may be prone to misinterpretation and will translate into less-accurate predictions, as the amount of information that can be digested effectively is limited (Kahneman 2003).

3. Data

In this section, we describe the comprehensive database we set up for the panel regressions reported in the next section. We first describe the dependent variables and then the independent variables. The observations are for a maximum of 73 countries from 1998 to (due to a data constraint) 2014, summing up to 17 years of 204 monthly observations per country and forecast variable. The panel exhibits missing values (unbalanced panel). For a full, detailed account of

[5]In another paper (Lustenberger and Rossi 2017b), we test the model by Morris and Shin (2002) on short-term and long-term interest rate forecasts.

the variables, we refer to the working paper version of this study (Lustenberger and Rossi 2017a).

3.1 Dependent Variables

The dependent variables are the absolute cross-sectional mean errors and the cross-sectional standard deviations of forecasts made by professional forecasters in predicting two financial variables, namely, short-term and long-term interest rates, and two macroeconomic variables, namely, consumer price index (CPI) inflation and the growth rates of real gross domestic product (GDP). All data are monthly from Consensus Economics. Consensus Economics groups countries in four geographic regions: Asia-Pacific (AP), Eastern Europe (EE), Latin America (LA), and "Western countries" (WE) (North America, Western Europe, Israel, Egypt, Nigeria, Saudi Arabia, and South Africa).

Each month, the survey participants for a particular country report their forecasts of short-term interest rates for 3 and 12 months ahead. They also report their view on long-term interest rates, i.e., the yields on their country's 10-year government debt, also 3 and 12 months ahead. Forecasts for CPI inflation and the growth rate of real GDP are also reported on a monthly basis but refer to the end of the current and the following year. Forecasts are provided by nongovernmental entities (independent or research institutes affiliated with universities) and economic consulting firms. The majority are financial institutions, varying from domestic and regional commercial banks to global investment banks.

We compare forecasts with realized short-term and long-term interest rates, as well as end-of-year consumer price indexes (where CPI was not available, we chose the GDP deflator) and growth in real GDP. The data are from Reuters Eikon, Bloomberg, IMF International Financial Statistics, and the World Bank database.

3.2 Independent Variables

All independent variables are observed at a time when the forecasts are published by Consensus Economics (on that day or the day before). We call this point in time the "forecast formation date."

Speeches. We construct a new explicit measure of communication consisting of central bank speeches. To this end, we compiled a variable made up of central bank speeches as collected by the Bank for International Settlements. For each central bank reporting their speeches to the BIS, we counted the number given in the month preceding the forecast.

Table 1 exhibits the total number of speeches per country and their monthly average divided by four geographic areas. As can be seen, most speeches are given by central banks in "Western countries" (WE), above all by the Federal Reserve (1,386) and Japan (453). Indian central bankers, grouped with the Asia-Pacific countries (AP), delivered the second-highest number of speeches (648). Figure 1 illustrates how communication activities by central banks have intensified over time. The number of speeches has steadily increased from approximately 150 in 1998 to nearly 900 in 2013 and 2014. Importantly, this reflects not only more communication activities but also a higher number of central banks reporting their speeches to the BIS since 2000 ("New Central Banks"). These are shown by the light gray columns (light blue in online version).

There is no potential endogeneity problem associated with our communication measure. Central bank speeches are announced months in advance. Therefore, the number of speeches is fixed, making our communication proxy a well-defined exogenous variable. We cannot exclude that some of the speeches were the result of unexpected events that the central bank considered important enough to justify intervention. However, for the bulk of the speeches, this is very unlikely. While the number of speeches is fixed, their content may take the economic situation into account. Hence, content analysis would raise problems of endogeneity. For this reason, we use the number of speeches as our communication variable and refrain from content analysis.

Transparency. We employ the most comprehensive measure of central bank transparency in terms of country and time coverage based on an extension of the multiple-dimensional transparency index of Eijffinger and Geraats (2006) presented by Dincer and Eichengreen (2014). The index includes five dimensions: political transparency, economic transparency, procedural transparency, policy transparency, and operational transparency. Each dimension covers three questions—for instance, Does the central bank have an explicit objective? Does the central bank disclose macroeconomic models? Does the central bank publish voting records of their board

Table 1. Central Bank Speeches by Country from 1998 to 2014

WE (24)			AP (15)			EE (24)			LA (10)		
	#	Mean		#	Mean		#	Mean		#	Mean
USA	1,386	6.79	AUS	280	1.37	CZE	35	0.17	ARG	31	0.15
JPN	453	2.22	CHN	93	0.46	HUN	11	0.05	BRA	10	0.05
DEU*	361	1.88	HKG	171	0.84	POL	20	0.10	CHL	73	0.36
FRA*	146	0.76	IND	648	3.18	RUS	3	0.01	MEX	40	0.20
GBR	373	1.83	IDN	36	0.18	TUR	83	0.41	VEN	0	0.00
ITA*	148	0.77	MYS	293	1.44	BGR	12	0.06	COL	5	0.02
CAN	344	1.69	NZL	110	0.54	HRV	4	0.02	PER	0	0.00
NLD*	88	0.46	PHL	204	1.00	EST**	19	0.09	URY	1	0.00
NOR	210	1.03	SGP	148	0.73	LVA**	8	0.04	SLV	0	0.00
ESP*	138	0.72	KOR	59	0.29	LTU**	0	0.00	GTM	1	0.00
SWE	410	2.01	TWN	0	0.00	ROU	24	0.12			
CHE	281	1.38	THA	175	0.86	SVK**	3	0.01			
AUT*	60	0.31	BGD	0	0.00	SVN**	2	0.01			
BEL*	31	0.16	PAK	108	0.53	UKR	0	0.00			
DNK	69	0.34	LKA	57	0.28	ALB	190	0.94			
FIN*	80	0.42				ARM	1	0.00			
GRC**	54	0.26				AZE	0	0.00			
IRL*	126	0.66				BLR	0	0.00			
PRT*	21	0.11				BIH	6	0.03			

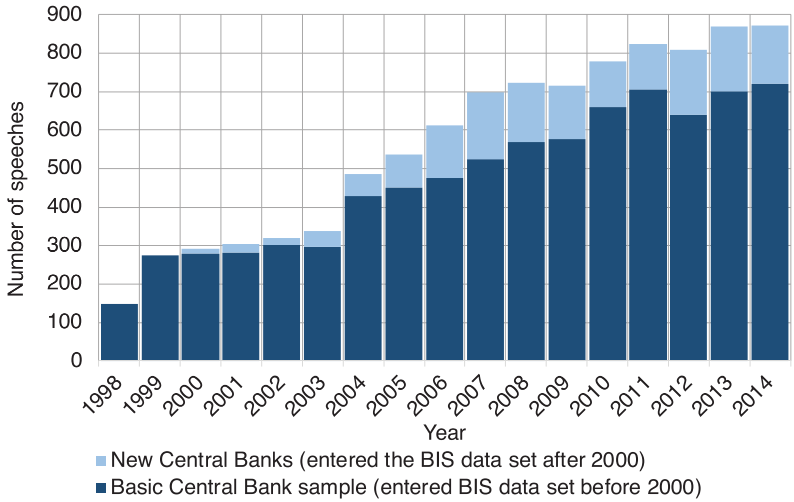
(*continued*)

Table 1. (Continued)

WE (24)			AP (15)			EE (24)			LA (10)		
	#	Mean		#	Mean		#	Mean		#	Mean
EGY	0	0.00				CYP**	3	0.01			
ISR	75	0.37				GEO	0	0.00			
NGA	29	0.14				KAZ	0	0.00			
SAU	28	0.14				MKD	39	0.19			
ZAF	233	1.14				MDA	0	0.00			
ECB*	1,386	7.22									
LUX*	37	0.19									
MLT**	19	0.23									

Notes: The table shows the total number of speeches (#) and the average number of speeches per month (mean) for a country. WE denotes "Western countries," AP Asia-Pacific countries, EE Eastern European countries, and LA Latin American countries. The number of countries in the set is given in parentheses. The euro was introduced in 1999. The ECB counts speeches given by its Executive Board members. In addition, we count all speeches given by members of the Eurosystem from 1999 onward, marked with *. Therefore, the total number of speeches used in the variable speech for AUT, BEL, FIN, FRA, DEU, IRL, ITA, NLD, PRT, and ESP is # 2,712 (mean 14.26). Countries which entered the Eurosystem after 1999 are marked with **. For these countries, all Eurosystem speeches are included following their adherence to the Eurosystem. GRC entered in 2001 (# 2,554 and mean 12.52), SVN in 2007 (# 1,961 and mean 9.66), CYP in 2008 (# 1,780 and mean 8.77), SVK in 2009 (# 1,533 and mean 7.55), EST in 2011 (# 1,086 and mean 5.35), and LVA since 2014 (# 287 and mean 1.41). LTU is not a member of the Eurosystem in our sample, for it joined in 2015. No forecasts for LUX and MLT (which entered the Eurosystem in 2008) are available. We use LUX and MLT for completeness of the total number of speeches in the Eurosystem.

Figure 1. Total Speeches per Year



Notes: Figure 1 displays the number of speeches given by representatives of central banks per year in the countries included in our sample. The basic central bank sample includes USA, JPN, DEU, FRA, GBR, ITA, CAN, NLD, NOR, ESP, SWE, CHE, AUT, BEL, GRC, ZAF, AUS, CHN, HKG, IND, IDN, MYS, NZL, SGP, KOR, TUR, LVA, BRA, and the ECB. For these central banks, there is at least one speech in 1998, in 1999, or in both.

members? The index has annual observations from 1998 to 2010 for 120 central banks.[6] A score for each central bank between 0 (minimum transparency) and 15 (maximum transparency) can be obtained. We measure transparency by the updated values of Dincer and Eichengreen, which extends the observations reported in Dincer and Eichengreen (2014) by four more years, until 2014.

Politico-Institutional Framework of Central Bank Independence. We capture the politico-institutional framework at central banks with the actual turnover of the central bank's governor in a year, as described by Dreher, Sturm, and de Haan (2010). What does the turnover rate stand for in our setup? The literature on central bank independence uses this variable as an indicator for central

[6] As documented by Dincer and Eichengreen (2014), central banks in countries with higher per capita income, deeper financial markets, more-open economies, and stronger political institutions are more likely to be more transparent than others.

bank independence, but on average over time. If a central bank has more governors in a given period, it is presumably less independent. In our application, the interpretation is different. The variable is equal to one when the central bank governor changes. Accordingly, uncertainty about future central bank behavior might easily be greater, leading to more inaccurate and dispersed forecasts. However, this is an effect of uncertainty about the person and is unrelated to the independence of the central bank. If anything, the personality of the governor should be more important in an independent central bank. This implies that in such a central bank, forecast dispersion is more likely than in a central bank where the governor changes but is known to take instructions from the government.[7]

Uncertainty Measures. We use the Chicago Board Options Exchange Volatility Index (VIX) as an uncertainty measure. The VIX is observed the day before forecasts are made. Broadly speaking, the VIX may also account for the business (interest rate) cycle, which is typically neglected in the related literature.

ZLB Dummy. We created a dummy variable for the period during which the zero lower bound (ZLB) on nominal interest rates was binding. We set the binding constraint at an interest rate level below 0.5 percent. It is only employed in the regressions of interest rate forecasts.

IT Dummy. We created a dummy that is equal to one if a central bank pursues an inflation targeting policy, and zero otherwise.

FG Dummy. We also accounted for central banks pursuing a forward guidance policy. We created a dummy that is equal to one during a period when forward guidance was pursued, and zero otherwise. We set the dummy based on Charbonneau and Rennison (2015). Accordingly, the following central banks made use of forward guidance: Japan from April 1999 to July 2000 and from October 2010 to March 2013, the United States from December 2008 to December 2014 (end of our sample period), Canada from April 2009 to April 2010, Sweden from April 2009 to July 2010 and February 2013 to December 2014, the European Central Bank from July 2013 to December 2014, and GBR from August 2013 to December 2014. All euro-area countries are equated with the European Central Bank

[7]We thank an anonymous referee for pointing this out to us.

Table 2. Summary Statistics for Forecast Errors (FE), Cross-Sectional Standard Deviation of Forecasts (Std.), Central Bank Transparency, Speeches, and Additional Variables

Variable	Mean	Std. Dev.	Min.	Max.	N
Short-Term Interest Rate (Abs. FE)	1.093	2.536	0	67.293	11,690
Long-Term Interest Rate (Abs. FE)	0.652	0.591	0	8.467	7,158
Inflation (Abs. FE, Current Year)	1.099	3.641	0	158.379	12,307
Inflation (Abs. FE, Next Year)	2.214	5.871	0	236.979	12,286
Growth (Abs. FE, Current Year)	1.3	1.394	0	12.557	11,773
Growth (Abs. FE, Next Year)	2.14	2.468	0.001	24.112	11,686
Short-Term Interest Rate (Std.)	0.52	1.281	0	41.598	11,806
Long-Term Interest Rate (Std.)	0.379	0.657	0	15.909	7,718
Inflation (Std., Current Year)	0.572	1.719	0.019	80.525	7,459
Inflation (Std., Next Year)	0.839	2.334	0	108.383	7,447
Growth (Std., Current Year)	0.453	0.391	0.026	8.653	7,459
Growth (Std., Next Year)	0.594	0.41	0.054	7.722	7,447
Central Bank Transparency					
Overall	7.7	3.29	0	15	24,593
"Western Countries" (WE)	9.4	2.97	1	15	9,456
Asia-Pacific (AP)	6.7	3.27	0	14	5,757
Eastern Europe (EE)	7.2	3.07	1	14.5	6,376
Latin America (LA)	5.9	2.14	1	9	3,004
Speech	3.51	7.374	0	50	24,593
Turnover	0.16	0.37	0	3	24,593
VIX	21.33	9.21	10.02	70.33	24,593
Inflation Targeting	0.32	0.466	0	1	24,593
Forward Guidance	0.04	0.195	0	1	24,593

Notes: The table provides summary statistics for absolute cross-sectional mean forecast errors (Abs. FE) and cross-sectional standard deviations (Std.) for the two financial variables (short-term and long-term interest rates) and the two macroeconomic variables (inflation and growth, with current-year forecasts and next-year forecasts). The table also exhibits the variation in central bank transparency in the four subsets of geographic regions ("Western," Asia-Pacific, Eastern European, and Latin American countries). Speech is the number of speeches held by central banks per month. Turnover measures replacement of central bank governors. VIX is the Chicago Board Options Exchange Volatility Index. Inflation Targeting is a dummy for inflation-targeting policy, and Forward Guidance is a dummy for forward guidance as policy instrument.

(ECB) forward-guidance dummy for the corresponding months. LVA is given a value of 1 from January 2014 after joining the currency union.

Table 2 offers summary statistics for the variables of the benchmark regressions. In the first set, we report the statistics for the absolute forecast errors of the four dependent variables. The absolute cross-sectional mean forecast errors range from zero to a maximum of 237 percent for inflation forecasts. In the second set, we present the corresponding cross-sectional standard deviations, which range from a minimum of zero to a maximum of 108. The third set yields the details of the transparency index. It covers the whole (theoretical) range from zero to 15. In addition to the overall values, we report the values for each of the four geographic areas. The fourth set contains summaries of the other independent variables. The number of speeches (our communication variable) has a mean of 3.5 per month and a maximum of 50. For annual values of turnover of central bank governors, the minimum is zero and the maximum reaches 3. The VIX ranges from a minimum value of 10 to a maximum of 70, with an average of 21. In 11 percent of the total number of observations, the zero lower bound was effective. About 30 percent of our observations are in countries with inflation-targeting policies. Around 4 percent of observations coincide with forward-guidance periods.

4. Results

This section is divided into eight subsections. In subsection 4.1, we describe our benchmark model and compare it with previous papers. In subsection 4.2, we elaborate on our benchmark regression results, and in subsection 4.3 we highlight differences to the previous literature. Subsection 4.4 examines the impact of the level of transparency. Subsection 4.5 deals with the outcome from an inflation-targeting regime. Subsection 4.6 discusses the effects of forward guidance, subsection 4.7 analyzes regional effects, and subsection 4.8 compares the results since the financial crisis with those obtained before the crisis.

4.1 Benchmark Model

We begin with an explanation of our basic fixed-effects regression model. It is given by

$$Y_{i,h,t} = \alpha + \nu_i + \beta_{SP} \cdot \text{Speech}_{i,t} + \beta_{TI} \cdot \text{Transp.}_{i,t} + \beta_{TO} \cdot \text{Turnover}_{i,t} + \beta_{VIX} \cdot \text{VIX}_t + \beta_H \cdot H_h + \beta_T \cdot T_y + \beta_{ZLB} \cdot ZLB_{i,t} + \varepsilon_{i,h,t},$$

where i is the country, h is the forecast horizon, and t is a monthly time index.

Our left-hand-side variable $Y_{i,h,t}$ is either the logarithm of the absolute cross-sectional mean forecast error ($\log[|FE_{i,h,t}|]$) or the logarithm of the cross-sectional standard deviation of forecasts ($\log[\sigma_{i,h,t}]$) provided by Consensus Economics. Our forecast variables are short-term interest rates, long-term interest rates, the percent change per annum of the CPI, and the growth rate of real GDP.[8]

On the right-hand side, α is the intercept, and ν_i is the fixed effect for country i. $\text{Speech}_{i,t}$ captures the number of speeches held by central bank representatives of country i between $t-1$ and t. $\text{Transp.}_{i,t}$ denotes central bank transparency. $\text{Turnover}_{i,t}$ stands for the number of central bank governor turnovers, and VIX_t represents the volatility index. H_h is a horizon fixed effect for the forecast horizon, only used for CPI inflation and real GDP growth forecasts, T_y is a yearly fixed effect to capture a possible time trend, and $ZLB_{i,t}$ is a dummy for the zero lower bound (only used for short-term and long-term interest rates).

No potential endogeneity arises. All right-hand-side variables are taken at the point in time when a forecast is formed (or the day before). For instance, for the U.S. CPI inflation forecast formed on June 11, 2012 for the end-of-year CPI inflation, all right-hand-side variables are measured on June 11, 2012 (or June 10, 2012). For the number of speeches in June 2012, we counted all speeches held between May 15, 2012 (because the forecast day in the previous month was May 14, 2012) and June 11, 2012. As we described in subsection 3.2, speeches are announced months in advance, making our communication proxy a well-defined exogenous variable. The VIX is taken with its value on June 10, 2012 (the day before the forecast is formed). Hence, while higher uncertainty about stock prices on June 10, 2012 makes forecasts formed on June 11 more uncertain, a reverse effect is not possible. For this reason we do not treat survey dispersion as a measure of uncertainty as is commonly done in the literature. Turnover and transparency are also well-defined exogenous variables. It is highly unlikely that a major change in the

[8]Estimates for consumption growth and industrial production growth are available upon request.

management of the central bank, a turnover, materializes because of forecasters' performance. For transparency, it is impossible that one single, even extreme, forecast error realized in December 2012 motivates a central bank to change its transparency and communication policy in June 2012 (before the forecast error is observed).

Compared with previous research, we introduce three important new variables: the number of speeches, the turnover of the central bank's governor, and the zero-lower-bound constraint. These three variables extend the empirical literature on central bank transparency along three dimensions. They allow us to (i) make a clear distinction between transparency and communication; (ii) account for potential implications that uncertainty in a central bank's management can have on forecasts; and (iii) consider special constraints on monetary policy which arose due to the financial crisis.[9] Subsection 4.6 goes even further and extends our benchmark model with a variable for forward guidance.

We distinguish three measures of the dependent variables. The first is the absolute forecast error, the second is the cross-sectional standard deviation, and the third measure is the interquartile range. In the literature following Mankiw, Reis, and Wolfers (2003), the favorite measure of cross-sectional dispersion is the interquartile range of forecasts. Arguably, the advantage of this measure over the simple standard deviation is that it is insensitive to outliers, which might be important in the analysis of survey data.

A fundamental distinction from previous papers arises in the way we calculate the dependent variables. In order to avoid that the residuals follow a pattern, we take logs of the dependent variables.[10] Besides heteroskedasticity issues, there is a clear-cut lower bound to the value residuals assume in our setting. However, taking the log of the absolute cross-sectional mean forecast errors and the log of the cross-sectional standard deviation generates well-behaved residuals. Note that taking the log of the interquartile range does not eliminate the "truncated" pattern in the residuals. For this reason, we

[9] In addition, we employ the VIX instead of GARCH estimates as an uncertainty measure.

[10] We are grateful to Massimilano Marcellino for suggesting this data transformation.

only use the log of cross-sectional standard deviations of forecasts as a dispersion measure.

In addition, we use panel clustered standard errors, where we cluster around countries. Variance estimates using panel clustered standard errors are consistent, as shown by Stock and Watson (2008). The procedure has three advantages: it allows us to get rid of inconsistent variance estimates, to take heteroskedasticity into account,[11] and to correct for correlation in the forecast errors arising from overlapping forecast horizons.

4.2 Benchmark Regression Results

In this subsection, we discuss the results of the benchmark regressions. We pursue a conservative approach and execute two-sided tests for significance of the coefficients.[12] The results are summarized in table 3 for absolute forecast errors and for standard deviations.

The findings are sobering. Transparency does not improve the predictability of financial and macroeconomic variables. More effective is the influence of greater transparency on forecast dispersion by reducing the misalignment among forecasters of interest rates and, especially, inflation.

The main and most important novel contribution to the literature relates to communication, which, as discussed, is measured by the number of central bank speeches. Three results arise. First, communication exerts a much greater influence on private forecast performance than transparency. Second, intensive communication activities make it more difficult to forecast inflation and real GDP growth. More communication also increases the dispersion in forecasts of inflation and interest rates. Third, in terms of statistical

[11] As Santos Silva and Tenreyro (2006) pointed out, heteroskedastic residuals coming from a log-linearized model lead to biased estimates of the true model parameters. However, a visual inspection of the residuals of our benchmark model shows that at least the residuals under our log model for absolute forecast errors do not exhibit heteroskedasticity. In the log model for cross-sectional standard deviation, heteroskedasticity cannot be excluded by visual inspection, but seems to be rather a minor issue.

[12] A one-sided test would be appropriate if the estimated value departs from the reference value in only one direction. However, as summarized above, in theory, departures are possible in both directions.

Table 3. Absolute Cross-Sectional Mean Forecast Errors and Forecasts' Cross-Sectional Standard Deviations

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)							
	Inflation		**Growth**		**Short-Term Interest Rate**		**Long-Term Interest Rate**	
	CY	**NY**	**CY**	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.019^{***}	0.020^{***}	0.014^{***}	0.007	0.008^{*}	0.002	0.006	0.004
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Transp.	0.022	−0.010	0.031	0.055	−0.044	−0.021	−0.051	-0.103^{*}
	(0.03)	(0.04)	(0.03)	(0.04)	(0.06)	(0.06)	(0.04)	(0.05)
Turnover	0.258^{***}	0.073	0.046	0.085	0.126	0.281^{***}	0.032	0.016
	(0.08)	(0.08)	(0.07)	(0.09)	(0.10)	(0.09)	(0.09)	(0.10)
VIX	−0.003	0.002	0.002	-0.003^{**}	0.015^{***}	0.001	0.011^{***}	-0.013^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ZLB					-0.723^{***}	-0.736^{***}	0.068	0.233^{*}
					(0.16)	(0.12)	(0.11)	(0.12)
N	12,307	12,286	11,787	11,710	5,828	5,858	3,565	3,590
Countries	73	73	73	73	34	34	23	23
R2	0.21	0.15	0.12	0.14	0.21	0.23	0.07	0.19

(*continued*)

Table 3. (Continued)

	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)							
	Inflation		**Growth**		**Short-Term Interest Rate**		**Long-Term Interest Rate**	
	CY	**NY**	**CY**	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.012***	0.008**	−0.001	0.004*	0.013***	0.002	0.007**	0.005**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Transp.	−0.109**	−0.115***	−0.017	−0.023	−0.090*	−0.082**	−0.107**	−0.091**
	(0.04)	(0.04)	(0.03)	(0.02)	(0.05)	(0.04)	(0.05)	(0.04)
Turnover	0.125	0.094	0.061	0.048	0.165**	0.087	0.104**	0.056*
	(0.09)	(0.07)	(0.04)	(0.04)	(0.08)	(0.07)	(0.04)	(0.03)
VIX	0.001	0.006***	0.005***	0.008***	0.014***	0.005***	0.010***	0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ZLB					−0.349*	−0.207	0.118	0.144*
					(0.18)	(0.13)	(0.07)	(0.07)
N	7,471	7,471	7,471	7,471	5,895	5,916	3,855	3,859
Countries	45	45	45	45	34	34	23	23
R2	0.46	0.24	0.50	0.27	0.12	0.34	0.20	0.17

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of the log of absolute cross-sectional mean forecast errors and the log of cross-sectional standard deviations. Inflation and growth have forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. Short-term and long-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year (the intercept, fixed effects, and yearly dummies are not shown in the table).

significance, the effect of speeches on inflation forecasts is highest. We provide a discussion and interpretation of these results in section 6.

The further analysis sheds light on the effects of politico-institutional aspects of central banks as measured by the turnover rate. The results suggest that instability in the management of central banks (turnover variable) makes the inflation outlook for the current year more uncertain and also reduces the predictability of future interest rate actions of the central bank in the 12 months ahead. Moreover, interest rate forecasts at the three-month horizon become more heterogenous.

The VIX plays an important role in all regressions of forecast dispersion but one. The higher market uncertainty is, the wider the dispersion becomes. In the same vein, market uncertainty makes interest rate predictions at the three-month horizon less precise. By contrast, it increases the accuracy of next year's GDP and long-term interest rate forecasts.

Another result is that when the zero lower bound is reached, short-term interest rates become easier to predict. By contrast, the zero-lower-bound constraint has no effect on forecasts of long-term interest rates. The latter finding is in line with Jain and Sutherland (2018). Neither does the ZLB affect the dispersion of interest rate forecasts.

Finally, the fit of the regressions is higher for dispersion as a dependent variable. For this type of regression, the best fit is found for growth forecasts and inflation.

4.3 Differences to the Literature

As we pointed out in section 1, we depart from previous papers along several dimensions. In terms of estimation strategy, we take the log of the independent variables instead of their level. This leaves the results unaffected. Hence, while the log generates well-behaved residuals, it does not affect the results either in quantitative or qualitative terms. Detailed results are shown in the appendix in tables A.1–A.4.

By contrast, including the yearly dummy in the benchmark regression leads qualitatively to similar results, but the number of significant coefficients in the regressions on transparency decreases.

Thus, our use of a time dummy explains largely the difference to the findings of previous work that reports significant effects of transparency such as Naszodi et al. (2016). However, our results relating to communication are not affected by the time dummy, except for the accuracy of long-term interest rates, whose coefficient becomes insignificant.

Moreover, the absolute oil price change is often used in such regressions. Large changes in oil prices are associated with larger uncertainty. Since in most cases this variable was not significant, we refrain from including it in our benchmark regression.

4.4 Impact of Transparency Level (Breakpoint 10)

Are there diminishing marginal returns from greater transparency? To answer this question, we analyze whether the worsening in the accuracy of forecasts arising from more communication shown above also holds at central banks whose transparency level is below 10, using a corresponding dummy. This threshold corresponds to the upper third of possible levels. As can be seen in table 4, except for short-term growth forecasts, more communication has similar effects on forecast errors irrespective of levels of transparency. Hence, the impaired quality of forecasts that arises from more communication observed in the whole sample cannot be attributed to distinct degrees of transparency. In this sense, we hardly find evidence of diminishing marginal returns from more transparency.

4.5 Inflation Targeting

A monetary policy strategy widely deemed to increase the transparency of policymaking is inflation targeting (IT), and some attention has been devoted to its effect on forecast performance. Cecchetti and Hakkio (2009) estimate how it affects the dispersion of private-sector forecasts of inflation. Using a panel data set that includes 15 countries over 20 years, they find no convincing evidence that IT reduces forecast dispersion. The results reported by Crowe (2010) for 11 countries suggest that IT improved the inflation forecasts for those whose initial forecast accuracy was worst, without harming the best forecasters.

We add to the evidence by assessing whether and how the introduction of IT has had any repercussion on the quality and standard

Table 4. Interaction Dummy for Transparency Below 10 and Absolute Cross-Sectional Mean Forecast Errors

	Inflation		Growth		Short-Term Interest Rate		Long-Term Interest Rate	
	CY	**NY**	**CY**	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.019***	0.021***	0.011**	0.006	0.009**	0.001	0.006*	0.004
Speech * Lower	−0.025	−0.010	0.039**	0.006	−0.000	0.017	−0.039	−0.037
Transp.	0.021	−0.007	−0.009	0.047	−0.026	−0.049	−0.050	−0.185**
Lower	−0.003	0.034	−0.363**	−0.067	0.125	−0.213	0.064	−0.243

Notes: Country fixed-effects panel regression with panel clustered standard errors. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows summary results for a fixed-effects panel regression of the log of absolute cross-sectional mean forecast errors using an interaction dummy which is one if transparency is below 10 (Lower). Inflation and growth have forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. Short-term and long-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year, the VIX, and the turnover. Moreover, we include the ZLB dummy in interest rate regressions (these variables, the intercept, fixed effects, yearly dummies, and horizon dummies are not shown in the table).

deviation of forecasts of inflation, as well as of short-term and long-term interest rates. For this, we added a dummy for the presence of IT, which interacts with our communication and transparency variables, to the regressors of the benchmark analysis ($\mathrm{IT}_{i,t}$). The regression equation is now

$$\begin{aligned} Y_{i,h,t} = \alpha + \nu_i &+ \beta_{SP} \cdot \mathrm{Speech}_{i,t} + \beta_{SPIT} \cdot \mathrm{Speech}_{i,t} \cdot \mathrm{IT}_{i,t} \\ &+ \beta_{TI} \cdot \mathrm{Transp.}_{i,t} + \beta_{TIIT} \cdot \mathrm{Transp.}_{i,t} \cdot \mathrm{IT}_{i,t} + \beta_{IT} \cdot \mathrm{IT}_{i,t} \\ &+ \beta_{TO} \cdot \mathrm{Turnover}_{i,t} + \beta_{VIX} \cdot \mathrm{VIX}_t + \beta_H \cdot H_h + \beta_T \cdot T_y \\ &+ \beta_{ZLB} \cdot ZLB_{i,t} + \varepsilon_{i,h,t}. \end{aligned}$$

Some interesting results emerge from this analysis. As can be inferred from table 5, communication and transparency of inflation targeters affect the quality of short-term interest rate forecasts in opposite ways. Whereas increased communication improves their accuracy, greater transparency leads to less-accurate forecasts. Moreover, inflation targeting weakens the forecast alignment effect from more transparency. By contrast, communication of inflation targeters does not affect forecasts' heterogeneity differently from non-inflation targeters. Overall, it seems that higher levels of transparency are especially critical in countries that pursue an inflation-targeting strategy.

4.6 Forward Guidance

In this subsection, we examine whether forward guidance has improved the predictability of interest rates. For this purpose, we add a dummy capturing forward guidance ($\mathrm{FG}_{i,t}$) to the list of benchmark regressors, which interacts with both the speech and the transparency variables. The fixed-effects regression model becomes

$$\begin{aligned} Y_{i,h,t} = \alpha + \nu_i &+ \beta_{SP} \cdot \mathrm{Speech}_{i,t} + \beta_{SPFG} \cdot \mathrm{Speech}_{i,t} \cdot \mathrm{FG}_{i,t} \\ &+ \beta_{TI} \cdot \mathrm{Transp.}_{i,t} + \beta_{TIFG} \cdot \mathrm{Transp.}_{i,t} \cdot \mathrm{FG}_{i,t} + \beta_{FG} \cdot \mathrm{FG}_{i,t} \\ &+ \beta_{TO} \cdot \mathrm{Turnover}_{i,t} + \beta_{VIX} \cdot \mathrm{VIX}_t + \beta_h \cdot H_h + \beta_T \cdot T_y \\ &+ \beta_{ZLB} \cdot ZLB_{i,t} + \varepsilon_{i,h,t}. \end{aligned}$$

As outlined in table 6, the combination of increased communication with forward guidance has effects on the quality of interest

Table 5. Inflation Targeting: Absolute Cross-Sectional Mean Forecast Errors and Forecasts' Cross-Sectional Standard Deviations

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)					
	Inflation		Short-Term Interest Rate		Long-Term Interest Rate	
	CY	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.019***	0.021***	0.016***	0.007**	0.005	0.004
Speech * IT	0.002	0.021	−0.079***	−0.063***	0.002	−0.011
Transp.	0.027	−0.018	−0.141**	−0.080	−0.013	−0.092
Transp. * IT	0.051	0.096*	0.208***	0.166**	−0.053	−0.014
IT	−0.823***	−1.270***	−1.804***	−1.697***	0.508	0.153
	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)					
	Inflation		**Short-Term Interest Rate**		**Long-Term Interest Rate**	
	CY	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.014***	0.009***	0.015***	0.001	0.008**	0.005**
Speech * IT	0.002	0.010	−0.000	0.014	−0.010	−0.003
Transp.	−0.132***	−0.133***	−0.107**	−0.062	−0.111**	−0.086**
Transp. * IT	0.104**	0.098**	0.104**	0.048	0.096*	0.080**
IT	−1.113***	−1.113***	−1.295***	−0.957**	−1.145**	−1.027**

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows summary results for a fixed-effects panel regression of the log of forecasts' cross-sectional absolute mean errors and standard deviations including a dummy for inflation targeting. Inflation has forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. Short-term and long-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year, the VIX, and the turnover. Moreover, we include the ZLB dummy in interest rate regressions (these variables, the intercept, fixed effects, yearly dummies, and horizon dummies are not shown in the table).

Table 6. Forward Guidance: Absolute Cross-Sectional Mean Forecast Errors and Forecasts' Cross-Sectional Standard Deviations

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)				Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)			
	Short-Term Interest Rate		Long-Term Interest Rate		Short-Term Interest Rate		Long-Term Interest Rate	
	M03	**M12**	**M03**	**M12**	**M03**	**M12**	**M03**	**M12**
Speech	0.009*	−0.001	0.002	−0.002	0.015***	0.008**	0.007**	0.005***
Speech * FG	−0.002	0.009	0.018**	0.017**	0.002	−0.021***	0.002	−0.000
Transp.	−0.042	−0.020	−0.048	−0.091*	−0.090*	−0.082**	−0.107**	−0.091**
Transp. * FG	0.085	0.019	0.065**	0.002	0.069	0.006	−0.009	0.003
FG	−0.837	−0.206	−0.957***	0.146	−1.121**	−0.011	0.051	−0.080

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows summary results for a fixed-effects panel regression of the log of absolute cross-sectional mean forecast errors and standard deviations including a dummy for forward-guidance periods. Short-term and long-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year, the VIX, and the turnover. Moreover, we include the ZLB dummy in interest rate regressions (these variables, the intercept, fixed effects, and yearly dummies are not shown in the table).

rate forecasts that differ from those on their homogeneity. On the one hand, more communication increases the inaccuracy of long-term interest rate forecasts. The same effect is caused by more transparency (three months ahead). On the other hand, intensified communication reduces the heterogeneity of short-term interest rate forecasts one year ahead. Summarizing, more communication and transparency under forward guidance makes forecasts of long-term interest rates less accurate.

4.7 Results Related to Geographic Regions

We also ran our benchmark regression for the four regions defined by Consensus Economics. The regions are "Western countries" (WE), Asia-Pacific (AP), Eastern Europe (EE), and Latin America (LA).

Overall, we find region-specific results. More-frequent communication worsens CPI inflation forecast accuracy in "Western countries" and in the Asia-Pacific region. On the other hand, higher transparency makes it more difficult to forecast inflation in Latin America, but it improves the forecast accuracy of both short-term rates at the 3-month forecast horizon in the Asia-Pacific area and long-term interest rates in "Western countries" at the 12-month horizon. For real GDP growth forecasts, we do not find any particular regional effect, neither for communication nor for transparency.

As far as forecasts' cross-sectional standard deviation is concerned, we do not observe any regional effect related to transparency. By contrast, for communication we find that in Latin American countries more-frequent communication decreases real GDP growth forecasts' dispersion. However, more-frequent communication increases dispersion of short-term interest rate forecasts at the 3-month horizon in "Western countries" and long-term interest rate forecasts at the 12-month forecast horizon in Eastern European countries.

4.8 Financial Crisis of 2007–08

Has the financial crisis influenced the quality and alignment of forecasts? In order to analyze this important question, we constructed a dummy for the financial crisis, starting in August 2007. The results are in table 7. According to it, the effects of transparency and

Table 7. Absolute Cross-Sectional Mean Forecast Errors, Forecasts' Cross-Sectional Standard Deviations, and the Financial Crisis of 2007–08

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)							
	Inflation		Growth		Short-Term Interest Rate		Long-Term Interest Rate	
	CY	NY	CY	NY	M03	M12	M03	M12
Speech	−0.001	−0.023**	0.005	−0.004	−0.022***	−0.008	0.011*	0.003
Speech * Crisis	0.016	0.036***	0.007	0.006	0.029***	0.008	−0.004	0.003
Transp.	0.011	−0.027	0.026	0.035	−0.057	−0.049	−0.056	−0.108**
Transp. * Crisis	0.035*	0.060***	0.016	0.048*	0.047	0.064**	0.065***	0.128***
Crisis	−0.138	−0.586***	−0.147	−0.267	−0.562	−0.410	−0.763***	−0.458
	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)							
	Inflation		**Growth**		**Short-Term Interest Rate**		**Long-Term Interest Rate**	
	CY	**NY**	**CY**	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.002	0.001	−0.004	−0.007*	−0.006	−0.008	−0.008**	0.003
Speech * Crisis	0.007	0.005	0.002	0.010**	0.017*	0.008	0.017***	0.003
Transp.	−0.124***	−0.128***	−0.024	−0.029	−0.108**	−0.049	−0.095**	−0.108**
Transp. * Crisis	0.042***	0.037**	0.020**	0.022***	0.050*	0.064**	0.046***	0.128***
Crisis	−0.346**	−0.209	−0.186**	−0.313***	−0.527*	−0.410	−0.682***	−0.458

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of the log of absolute cross-sectional mean forecast errors and the log of cross-sectional standard deviations using an interaction dummy which is one after July 2007—the beginning of the financial crisis. Inflation and growth have forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. Short-term and long-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year, the VIX, and the turnover. Moreover, we include the ZLB dummy in interest rate regressions (these variables, the intercept, fixed effects, yearly dummies, and horizon dummies are not shown in the table).

communication have changed since the financial crisis. Before the crisis, more-intense communication reduced the errors in forecasting inflation (for the next year) and short-term interest rates (for the next three months). Since the crisis, more communication has worsened the quality of these forecasts. Moreover, more communication has increased the dispersion of real GDP forecasts for the next year and long-term interest rate forecasts for three months ahead.

Related to transparency, the evidence is even more compelling. Since the financial crisis, greater transparency has impaired the quality of forecasts, especially those relating to inflation and short-term and long-term interest rates. In terms of forecast dispersion, while more transparency had a dispersion-decreasing effect, since the financial crisis this effect has decreased or has even been offset.

Overall, intensified communication and transparency since the financial crisis made it unambiguously more difficult for private forecasters to predict macro and financial variables compared with the pre-crisis period.

5. Robustness

We redid the estimations with a variety of alternative regressors. In sum, the benchmark results remained valid. We performed the following robustness checks.[13]

- *Transparency Subindexes:* Neuenkirch (2013) employs the overall transparency index by Eijffinger and Geraats (2006) and the five subindexes until 2009 for money market forecasts in 25 emerging market countries. The result is that all subindexes improve market expectations, with political transparency having the largest effect. We replace the overall transparency index by its five subindexes as regressors: political transparency, economic transparency, procedural transparency, policy transparency, and operational transparency.

  The results are heterogeneous. At odds with Neuenkirch (2013), we find no effect of any transparency subindex on short-term interest rates' forecast accuracy. By contrast, operational transparency has some beneficial effects in terms

[13]Estimates of all robustness checks are available upon request.

of accuracy and homogeneity of forecasts. In particular, it improves the accuracy of real GDP growth forecasts for next year and decreases heterogeneity of inflation and interest rate forecasts. Increasing political transparency produces mixed effects. On the one hand, it aligns short-term interest rate forecasts. On the other hand, it worsens the accuracy of real GDP forecasts for next year. Finally, more policy transparency worsens outcomes for long-term interest rate forecasts at the three-month horizon.

- *Excluding Euro-Area Countries:* So far, we have used all forecasts from the euro-area countries. To eliminate a potential overweight of the ECB—for instance, we have 17 euro-area countries in the calculation of absolute CPI forecast errors—we excluded all euro-area countries except for Germany[14] and redid the regression for forecast errors and cross-section standard deviations.

  Our benchmark results are confirmed. The only change compared with the benchmark is that transparency has a weaker (alignment-enhancing) effect on the standard deviations of short-term and long-term interest rate forecasts.
- *Uncertainty Measures:* We also checked whether the results remain unaltered if we replaced the VIX with the uncertainty measures from Jurado, Ludvigson, and Ng (2015) and the 30-day return volatility of country MSCI stock market indexes.

  Evidence obtained in the benchmark analysis remains unchanged for the measure by Jurado, Ludvigson, and Ng (2015). Transparency has no effect on forecast precision, but it does reduce forecast dispersion. By contrast, communication worsens the quality of forecasts and increases their dispersion. The effect of uncertainty itself on forecast errors is, contrary to the VIX utilized in the benchmark, significant for inflation forecasts. Contrary to the VIX, this uncertainty measure turns out to exert an insignificant effect on the standard deviation of inflation and GDP growth forecasts, for which the VIX yielded highly significant results. In addition, replacing the VIX with the 30-day return volatility of country MSCI

[14] For instance, Middeldorp (2011) only used observations from Germany in his analysis.

stock market indexes does not change the results related to transparency and communication either.

- *Only Basic Central Banks Sample for Speeches:*[15] As explained in subsection 3.2, the number of central banks reporting their speeches to the BIS increased over time. As a further robustness check, we repeated the benchmark analysis but limited it to the group of central banks, the majority, which has reported to the BIS on a regular basis since 1998. While the evidence of communication does not change, transparency, which in the benchmark analysis significantly reduces forecasts' misalignments, has no effect anymore. Therefore, we conclude that the dispersion-reducing effect of increased transparency that we observe in the benchmark regression is due to the newly reporting central banks.
- *Only Months with Speeches:* In the benchmark regressions we included observations in the months when actually no speech was delivered. As a robustness check, we eliminated these observations from the sample. The general message did not alter.[16]
- *Seasonality of Speeches:* We also ran our benchmark regression excluding July, August, and December since many countries have holidays during these months. Results do not change.
- *Exclusion of Outliers:* We excluded the countries with the largest outliers in terms of forecast errors and took into account the number of forecasters. The results do not change.[17]
- *Revised GDP Figures:* Finally, we found that real GDP growth forecasts react only a little to transparency and communication. However, as is well known, realized GDP figures are regularly revised. To account for this fact, we produced a proxy for GDP growth figures in real time to calculate absolute forecast errors. Our proxy is the December real GDP growth

[15]See figure 1.

[16]The coefficient of speech in the benchmark regression with absolute cross-sectional mean forecast errors as the dependent variable turned negative for interest rate forecasts but was largely insignificant.

[17]For CPI inflation forecasts, the largest absolute forecast errors (greater than 25 percentage points) are in ARG, BGR, BLR, MDA, RUS, TUR, UKR, and VEN. For interest rate forecasts, the outliers are ARG, IDN, and VEN.

forecasts for the current year (i.e., about three weeks' forecast horizon) as the nearest value to the effectively realized GDP growth figure. We reran the benchmark regression for absolute forecast errors using forecast errors calculated with this proxy. Overall, the main message concerning communication and transparency remains unchanged, but its significance decreased.

6. Policy Implications

How do we read the results found in this paper in terms of policy implications? When it comes to transparency, the policy implications are not clear-cut. If the policy objective is to get forecasters to provide more-precise forecasts, our results suggest that transparency is not an adequate tool to achieve it. If the objective is to align individual forecasts, then the general normative implication seems to be an increase in transparency. However, this result seems to be attributable to a selection of central banks that have more recently reported their speeches to the BIS.

In relation to speeches, our empirical results suggest that more-frequent communication increases the uncertainty of their recipients and/or that central banks communicate less precisely. This implies that in order to improve the quality of forecasts of variables that are central to monetary policymaking and align them among professional forecasters, central banks ought to speak less often.

The important question, then, is whether less-precise communication is an unintended effect of too much talk or whether it is a deliberate choice. The latter is not unrealistic. A case in point is particularly uncertain times, when the central bank may want to convey to the markets an increased uncertainty underlying its own forecasts or to dampen market participants' risk-taking behavior.

However, the second explanation is more realistic, for it is difficult to imagine that a central bank deliberately chooses to keep raising the uncertainty in the markets over time. The reason, as noted by Blinder (2007), is that the policy-effectiveness argument for central bank transparency boils down to teaching the markets to "think like the central bank." Doing so will enable the central bank to manage expectations of future monetary policy better and, in particular, to keep them in line with its own thinking.

According to Blinder (2004), there is no one "right way" to communicate. The most-appropriate forms of central bank communication with the public, the government, and the markets depend on the nature of the monetary policy committee. One potential disadvantage that is particularly relevant for an individualistic committee is that it may speak with too many voices, creating cacophony. Blinder (2007) argued that individualistic committees seemed to have coped with their potential cacophony problem.

Or have they? In a recent publication, Blinder (2018) seems to have a more nuanced view on this issue, predicting that the cacophony problem on monetary policy committees will not go away soon. The empirical evidence in this paper suggests that the number of speeches has created confusion rather than clarity.

To corroborate this conjecture, we regressed on a variable relating to central bank independence in lieu of (in)stability.[18] Both a higher turnover rate and a higher CBI index point qualitatively in the same direction. Both suggest a negative effect on the quality of forecasts, although they measure two distinct dimensions of the politico-institutional framework of central banks, mirrored in a correlation coefficient of –0.0026. The evidence seems reasonable for the turnover variable. More difficult is rationalizing the evidence associated with the CBI variable. After all, the central bank independence/credibility literature points to increased credibility arising from greater independence with potential favorable knock-on effects on the variance and predictability of variables. However, our CBI variable covers several dimensions of a central bank's structure. One possible dimension is consistent with the observation made by Blinder (2004) that central bank independence promotes the switch to committee decisions rather than individual decisions. This switch may have raised the risk that the central bank speaks with a cacophony of voices.

As an additional corroborating factor for this argumentation, we constructed a variable that captures the size of monetary policy

[18]Central bank independence (CBI) is proxied by the unweighted independence index constructed by Dincer and Eichengreen (2014). It runs from 0 (lowest independence) to 1 (highest independence) and ends in 2010. Benchmark model estimates replacing turnover with CBI are available upon request.

committees.[19] The correlation of committee size with the number of speeches is +0.74. This suggests that the number of speeches increases in committees' size. Moreover, the correlation of committee size with the CBI variable is +0.51. Hence, although causality may run in both directions, these correlations may suggest that the more independent central banks are, the larger the size of the monetary policy committees, and the larger the monetary policy committees are, the more speeches they deliver. This may be one explanation for the rise in the number of speeches over the years that underlie the rise of potential cacophony.

To sum up, we find that the more speeches central banks give, the greater the confusion this creates among forecasters. Whether this arises from too many voices about monetary policy issues or from too many topics not directly related to monetary policy decisions, such as climate change, education, or inequality, is an important avenue for future research. This is particularly important given our evidence that more communication and transparency have worsened the quality of forecasts and increased their misalignment especially after the financial crisis.

7. Conclusions

By increasing market participants' ability to predict future policy actions, transparency is expected to increase monetary policy effectiveness. Anticipation of the central bank's actions results in a smoother operation in the first steps of the transmission mechanism between policy actions and economic activity and inflation. The question therefore is: Does central bank communication and transparency affect macroeconomic forecasts at all and in the intended way? We answer this question based on a large sample of countries for financial and macroeconomic variables important for monetary policymaking.

The answer is only partially affirmative. Particularly, we show that a clear distinction between central bank communication and transparency should be made. Our main finding is that more speeches worsen the accuracy and precision of financial and macroeconomic forecasts. This insight had been anticipated by Simon

[19] See Lustenberger and Rossi (2017a) for details.

(1971), for whom "a wealth of information creates a poverty of attention." It also lines up with the conclusions drawn by, for instance, Morris and Shin (2002), Kahneman (2003), Sims (2003), and Blinder (2004), for whom uncoordinated communication might actually lower, rather than raise, the signal-to-noise ratio and, in turn, hamper the operation of monetary policy. Stated differently, a "central bank that speaks with a cacophony of voices may, in effect, have no voice at all" (Blinder 2004, p. 57). Thus, speaking less may be beneficial for central banks that want to raise predictability and homogeneity among financial and macroeconomic forecasts. We provide some evidence that this may be particularly true for central banks whose transparency level is already high.

Our results also show that we should not expect too much from greater transparency either. We provide compelling evidence that, in general, central bank transparency is not an effective instrument to improve the accuracy of private forecasts.

At a more detailed level, our results suggest that intensified communication and transparency since the financial crisis made it unambiguously more difficult for private forecasters to predict macro and financial variables compared with the pre-crisis period.

Overall, it seems that higher levels of transparency are especially critical in countries that pursue an inflation-targeting strategy.

A related question is whether a forward-guidance policy as pursued after the crisis has borne fruit. The evidence is mixed. Forward guidance combined with intensified communication worsened the quality of long-term interest rate forecasts but reduced the dispersion of short-term rate forecasts. Increased transparency in a forward-guidance regime produced partly less accurate long-term interest rate forecasts.

Overall, the empirical evidence supports the view expressed by Cukierman (2009), who probes the limits of transparency in general. He argues that given the high degree of opacity in the past, it is highly likely that the move of central banks over the last 20 years towards openness to the public has improved matters. However, he also reminds us that since sufficiently high transparency is now in place and is part of the orthodoxy, the time has come to take a more realistic look at the limits of its feasibility and desirability. That said, we also provide evidence that more transparency contributes to aligning single forecasts with each other. From this perspective, transparency seems to provide the anchor by which agents' forecasting actions are coordinated. Thus, what seems to be important in the discussion about more or less central bank transparency is to make a clear distinction between its effect on forecast accuracy and its effect on forecast dispersion.

In a more general context, our results make a contribution to the theoretical literature on the social value of information. The general message arising from our empirical analysis is that more public information may make it more difficult for agents to deduce its content.

We also make a contribution in highlighting the importance of the stability and independence of central banks to the quality of private-sector forecasts. Our finding on this account is that a higher turnover of governors tends to reduce the precision of short-term interest rate and inflation forecasts. Greater central bank independence could worsen the quality of forecasts, perhaps by increasing the size of monetary policy committees that may lead to cacophony.

We have identified the communication of the central bank by the number of speeches. Future research could extent the analysis. According to Freedman and Laxton (2009), there are additional important means for the central bank's views to reach the public, such as monetary policy reports, inflation reports, press releases, and minutes. Ideally, all means of central bank communication should complement each other to get their message across most effectively. Otherwise, there is a risk of overcommunicating and transmitting different messages through the various channels. While some of these communication tools are captured by our transparency index, a more in-depth analysis may provide evidence of their effectiveness. This is the case in Jain and Sutherland (2018), who identify the communication by the central bank projections typically disseminated in monetary policy reports and inflation reports and find the projections to be important inputs to private forecasts.

Future research could also track the evolution of committee sizes and analyze the effects per speaker. Another extension could examine the content of speeches along various dimensions, such as their length, comprehensibility and keywords, and accounting for endogeneity issues. This would allow to tackle the question of whether cacophony arises from too many speeches about all sorts of issues not directly related to monetary policy or from a variety of voices on monetary policy issues.

Appendix. Differences to the Previous Literature

Table A.1. Inflation, Previous Literature, and Time Dummy

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)											
	CY						NY					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Transp.	−0.229***	−0.108***	0.024	−0.272***	−0.129***	0.022	−0.397**	−0.077***	−0.005	−0.435**	−0.091***	−0.010
	(0.07)	(0.02)	(0.03)	(0.08)	(0.02)	(0.03)	(0.15)	(0.03)	(0.04)	(0.17)	(0.03)	(0.04)
VIX	0.019***	0.010***	−0.002	0.014***	0.008***	−0.003	0.013	0.005**	0.002	0.012*	0.006***	0.002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
$\|\Delta \text{Oil}\|$	−0.030***	−0.008*	−0.006*				−0.001	0.014***	0.004*			
	(0.01)	(0.00)	(0.00)				(0.02)	(0.01)	(0.00)			
Speech				0.027***	0.017***	0.019***				0.045***	0.026***	0.020***
				(0.01)	(0.00)	(0.00)				(0.01)	(0.00)	(0.01)
Turnover				0.948	0.212**	0.258***				1.054*	−0.015	0.073
				(0.63)	(0.08)	(0.08)				(0.59)	(0.09)	(0.08)
Time Dummies	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
N	12,322	12,319	12,319	12,310	12,307	12,307	12,298	12,298	12,298	12,286	12,286	12,286
Countries	73	73	73	73	73	73	73	73	73	73	73	73
R2	0.03	0.17	0.20	0.04	0.18	0.21	0.01	0.02	0.14	0.02	0.03	0.15

(continued)

Table A.1. (Continued)

	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)											
	CY						NY					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Transp.	−0.218***	−0.124***	−0.109**	−0.235***	−0.132***	−0.109**	−0.308***	−0.127***	−0.115***	−0.327***	−0.132***	−0.115***
	(0.07)	(0.03)	(0.04)	(0.07)	(0.03)	(0.04)	(0.09)	(0.03)	(0.04)	(0.10)	(0.03)	(0.04)
VIX	0.011***	0.011***	0.001	0.009***	0.011***	0.001	0.016***	0.012***	0.005***	0.015***	0.012***	0.006***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\|\Delta \text{Oil}\|$	−0.012*	0.000	0.001				−0.010	0.002	0.005***			
	(0.01)	(0.00)	(0.00)				(0.01)	(0.00)	(0.00)			
Speech				0.017***	0.012***	0.012***				0.020***	0.008**	0.008**
				(0.01)	(0.00)	(0.00)				(0.01)	(0.00)	(0.00)
Turnover				0.599	0.115	0.125				0.474	0.084	0.094
				(0.45)	(0.09)	(0.09)				(0.33)	(0.07)	(0.07)
Time Dummies	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
N	7,483	7,483	7,483	7,471	7,471	7,471	7,483	7,483	7,483	7,471	7,471	7,471
Countries	45	45	45	45	45	45	45	45	45	45	45	45
R2	0.06	0.41	0.45	0.08	0.42	0.46	0.05	0.19	0.23	0.06	0.20	0.24

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of (the log of) absolute cross-sectional mean forecast errors and (the log of) cross-sectional standard deviations. Inflation has forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. If a dummy for each year is included (time dummies), it is indicated by "Yes" (the intercept, fixed effects, time dummies, and horizon dummies are not shown in the table). Regression (1) corresponds to the regression by Naszodi et al. (2016), while (2) uses the log of the left-hand-side variable. (3) adds time dummies to otherwise the same regression as (2) (the same is true for (7)–(9)). Regression (4) corresponds to our regression set up without log and time dummies, while regression (5) uses the log of the left-hand-side variable. Additionally, (6) includes time dummies (the same is true for (10)–(12)).

Table A.2. Growth, Previous Literature, and Time Dummy

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)											
	CY						NY					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Transp.	−0.075**	−0.080***	0.034	−0093***	−0.089***	0.031	−0.057	−0.028	0.056	−0.016	−0.011	0.055
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.06)	(0.02)	(0.04)	(0.05)	(0.02)	(0.04)
VIX	0.017***	0.014***	0.002	0.016***	0.014***	0.002	0.037***	0.019***	−0.003**	0.046***	0.022***	−0.003**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
|ΔOil|	−0.007	−0.005	−0.000				0.076***	0.027***	0.003			
	(0.00)	(0.00)	(0.00)				(0.01)	(0.00)	(0.00)			
Speech				0.013**	0.005	0.014***				0.018	0.003	0.007
				(0.01)	(0.00)	(0.00)				(0.02)	(0.01)	(0.01)
Turnover				0.009	0.005	0.046				0.018	0.021	0.085
				(0.08)	(0.07)	(0.07)				(0.20)	(0.09)	(0.09)
Time Dummies	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
N	11,800	11,799	11,799	11,788	11,787	11,787	11,722	11,722	11,722	11,710	11,710	11,710
Countries	73	73	73	73	73	73	73	73	73	73	73	73
R2	0.08	0.07	0.11	0.08	0.07	0.12	0.06	0.04	0.14	0.04	0.04	0.14

(*continued*)

Table A.2. (Continued)

	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)											
	CY						NY					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Transp.	−0.036**	−0.045**	−0.018	−0.041**	−0.052**	−0.017	−0.044***	−0.043**	−0.024	−0.049***	−0.051***	−0.023
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
VIX	0.008***	0.014***	0.005***	0.007***	0.013***	0.005***	0.010***	0.014***	0.008***	0.009***	0.014***	0.008***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\|\Delta Oil\|$	−0.005***	−0.009***	−0.001				−0.004***	−0.005**	−0.001			
	(0.00)	(0.00)	(0.00)				(0.00)	(0.00)	(0.00)			
Speech				0.001	−0.001	−0.001				0.002	0.005*	0.004*
				(0.00)	(0.00)	(0.00)				(0.00)	(0.00)	(0.00)
Turnover				0.041	0.026	0.061				0.070*	0.029	0.048
				(0.03)	(0.04)	(0.04)				(0.04)	(0.04)	(0.04)
Time Dummies	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
N	7,483	7,483	7,483	7,471	7,471	7,471	7,483	7,483	7,483	7,471	7,471	7,471
Countries	45	45	45	45	45	45	45	45	45	45	45	45
R2	0.20	0.37	0.49	0.19	0.36	0.50	0.14	0.15	0.26	0.14	0.16	0.27

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of (the log of) absolute cross-sectional mean forecast errors and (the log of) cross-sectional standard deviations. Growth has forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. If a dummy for each year is included (time dummies), it is indicated by "Yes" (the intercept, fixed effects, time dummies, and horizon dummies are not shown in the table). Regression (1) corresponds to the regression by Naszodi et al. (2016), while (2) uses the log of the left-hand-side variable. (3) adds time dummies to otherwise the same regression as (2) (the same is true for (7)–(9)). Regression (4) corresponds to our regression set up without log and time dummies, while regression (5) uses the log of the left-hand-side variable. Additionally, (6) includes time dummies (the same is true for (10)–(12)).

Table A.3. Short-Term Interest Rates, Previous Literature, and Time Dummy

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)											
	M03						M12					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Transp.	−0.369**	−0.259***	−0.037	−0.364**	−0.206***	−0.044	−0.417**	−0.184***	−0.020	0.402**	−0.130***	−0.021
	(0.15)	(0.06)	(0.07)	(0.16)	(0.05)	(0.06)	(0.17)	(0.05)	(0.06)	(0.17)	(0.05)	(0.06)
VIX	0.028***	0.029***	0.014***	0.026***	0.027***	0.015***	0.031***	0.021***	0.001	0.029***	0.019***	0.001
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
$\lvert\Delta\text{Oil}\rvert$	−0.002	0.014	0.015**				0.003	0.013*	0.002			
	(0.02)	(0.01)	(0.01)				(0.02)	(0.01)	(0.00)			
Speech				0.010	0.001	0.008*				0.012	−0.004	0.002
				(0.01)	(0.00)	(0.00)				(0.01)	(0.00)	(0.00)
Turnover				1.189	0.098	0.126				1.190	0.227**	0.281***
				(0.84)	(0.10)	(0.10)				(0.78)	(0.10)	(0.09)
ZLB				0.013	−1.085***	−0.723***				−0.225	−1.003***	−0.736***
				(0.18)	(0.18)	(0.16)				(0.20)	(0.11)	(0.12)
Time Dummies	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
N	5,875	5,840	5,840	5,863	5,828	5,828	5,875	5,870	5,870	5,863	5,858	5,858
Countries	34	34	34	34	34	34	34	34	34	34	34	34
R2	0.07	0.11	0.19	0.11	0.15	0.21	0.07	0.08	0.20	0.10	0.14	0.23

(*continued*)

Table A.3. (Continued)

	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)											
	M03						M12					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Transp.	−0.218**	−0.211***	−0.085*	−0.232**	−0.205***	−0.090*	−0.264**	−0.165***	−0.078*	−0.276**	−0.158***	−0.082**
	(0.09)	(0.03)	(0.05)	(0.08)	(0.04)	(0.05)	(0.11)	(0.03)	(0.04)	(0.11)	(0.03)	(0.04)
VIX	0.010***	0.025***	0.014***	0.009***	0.022***	0.014***	0.008***	0.014***	0.006***	0.008***	0.012***	0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\|\Delta\text{Oil}\|$	−0.010	−0.005	0.004				−0.007	−0.006**	−0.000			
	(0.01)	(0.00)	(0.00)				(0.01)	(0.00)	(0.00)			
Speech				0.004	0.006**	0.013***				0.002	−0.003	0.002
				(0.00)	(0.00)	(0.00)				(0.00)	(0.00)	(0.00)
Turnover				0.636	0.160**	0.165**				0.562	0.085	0.087
				(0.46)	(0.08)	(0.08)				(0.41)	(0.07)	(0.07)
ZLB				0.406	−0.584***	−0.349*				0.502	−0.394***	−0.207
				(0.32)	(0.16)	(0.18)				(0.42)	(0.13)	(0.13)
Time Dummies	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
N	5,934	5,907	5,907	5,922	5,895	5,895	5,932	5,928	5,928	5,920	5,916	5,916
Countries	34	34	34	34	34	34	34	34	34	34	34	34
R2	0.09	0.08	0.11	0.15	0.10	0.12	0.11	0.23	0.33	0.15	0.27	0.34

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of (the log of) absolute cross-sectional mean forecast errors and (the log of) cross-sectional standard deviations. Short-term interest rates have forecasts for 3 months (M03) and 12 months ahead (M12). If a dummy for each year is included (time dummies), it is indicated by "Yes" (the intercept, fixed effects, and time dummies are not shown in the table). Regression (1) corresponds to the regression by Naszodi et al. (2016), while (2) uses the log of the left-hand-side variable. (3) adds time dummies to otherwise the same regression as (2) (the same is true for (7)–(9)). Regression (4) corresponds to our regression set up without log and time dummies, while regression (5) uses the log of the left-hand-side variable. Additionally, (6) includes time dummies (the same is true for (10)–(12)).

Table A.4. Long-Term Interest Rates, Previous Literature, and Time Dummy

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)											
	M03						M12					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Transp.	0.001	−0.014	−0.060	−0.001	−0.033	−0.051	0.052**	0.056*	−0.112**	0.022	0.020	−0.103*
	(0.01)	(0.03)	(0.04)	(0.01)	(0.03)	(0.04)	(0.02)	(0.03)	(0.05)	(0.00)	(0.03)	(0.05)
VIX	0.004***	0.010***	0.009***	0.005***	0.012***	0.011***	−0.003*	−0.002	−0.013***	−0.002	−0.000	−0.013***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\|\Delta Oil\|$	0.007***	0.007	0.012**				−0.004	0.002	0.001			
	(0.00)	(0.00)	(0.00)				(0.00)	(0.00)	(0.00)			
Speech				0.003*	0.009**	0.006				0.008***	0.012***	0.004
				(0.00)	(0.00)	(0.00)				(0.00)	(0.00)	(0.00)
Turnover				0.020	−0.006	0.032				−0.004	0.021	0.016
				(0.04)	(0.09)	(0.09)				(0.07)	(0.12)	(0.10)
ZLB				0.090**	0.242**	0.068				0.290***	0.434***	0.233*
				(0.04)	(0.11)	(0.11)				(0.08)	(0.12)	(0.12)
Time Dummies	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
N	3,580	3,577	3,577	3,568	3,565	3,565	3,602	3,602	3,602	3,590	3,590	3,590
Countries	23	23	23	23	23	23	23	23	23	23	23	23
R2	0.02	0.01	0.07	0.02	0.02	0.07	0.01	0.00	0.18	0.04	0.03	0.19

(*continued*)

Table A.4. (Continued)

	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)											
	M03						M12					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Transp.	−0.127	−0.076	−0.121**	−0.139	−0.093*	−0.107**	−0.136	−0.082**	−0.101**	−0.145	−0.092**	−0.091**
	(0.11)	(0.05)	(0.05)	(0.11)	(0.05)	(0.05)	(0.10)	(0.04)	(0.04)	(0.10)	(0.04)	(0.04)
VIX	0.004***	0.013***	0.010***	0.005***	0.013***	0.010***	0.003***	0.005***	0.005***	0.003***	0.006***	0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
|ΔOil|	−0.000	−0.002	−0.001				0.001	0.001	0.002			
	(0.00)	(0.00)	(0.00)				(0.00)	(0.00)	(0.00)			
Speech				0.006*	0.010***	0.007**				0.005*	0.006***	0.005**
				(0.00)	(0.00)	(0.00)				(0.00)	(0.00)	(0.03)
Turnover				0.137	0.103**	0.104**				0.110	0.048	0.056*
				(0.11)	(0.05)	(0.04)				(0.09)	(0.03)	(0.03)
ZLB				0.124	0.155***	0.118				0.125	0.139**	0.144*
				(0.08)	(0.05)	(0.07)				(0.08)	(0.05)	(0.07)
Time Dummies	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
N	3,871	3,867	3,867	3,859	3,855	3,855	3,871	3,871	3,871	3,859	3,859	3,859
Countries	23	23	23	23	23	23	23	23	23	23	23	23
R2	0.07	0.12	0.18	0.08	0.15	0.20	0.09	0.08	0.15	0.10	0.10	0.17

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of (the log of) absolute cross-sectional mean forecast errors and (the log of) cross-sectional standard deviations. Long-term interest rates have forecasts for 3 months (M03) and 12 months (M12) ahead. If a dummy for each year is included (time dummies), it is indicated by "Yes" (the intercept, fixed effects, and time dummies are not shown in the table). Regression (1) corresponds to the regression by Naszodi et al. (2016), while (2) uses the log of the left-hand-side variable. (3) adds time dummies to otherwise the same regression as (2) (the same is true for (7)–(9)). Regression (4) corresponds to our regression set up without log and time dummies, while regression (5) uses the log of the left-hand-side variable. Additionally, (6) includes time dummies (the same is true for (10)–(12)).

References

Amato, J. D., S. Morris, and H. S. Shin. 2002. "Communication and Monetary Policy." *Oxford Review of Economic Policy* 18 (4): 495–503.

Blinder, A. S. 2004. *The Quiet Revolution: Central Banking Goes Modern*. Yale University Press.

———. 2007. "Monetary Policy by Committee: Why and How?" *European Journal of Political Economy* 23 (1): 106–23.

———. 2018. "Through a Crystal Ball Darkly: The Future of Monetary Policy Communication." *AEA Papers and Proceedings* 108 (May): 567–71.

Blinder, A. S., M. Ehrmann, J. de Haan, and D.-J. Jansen. 2017. "Necessity as the Mother of Invention: Monetary Policy After the Crisis." *Economic Policy* 32 (92): 707–55.

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.

Cecchetti, S. G., and C. Hakkio. 2009. "Inflation Targeting and Private Sector Forecasts." NBER Working Paper No. 15424.

Charbonneau, K., and L. Rennison. 2015. "Forward Guidance at the Effective Lower Bound: International Experience." Staff Discussion Paper No. 2015-15, Bank of Canada.

Chortareas, G., D. Stasavage, and G. Sterne. 2002. "Does It Pay to Be Transparent? International Evidence from Central Bank Forecasts." *Review* (Federal Reserve Bank of St. Louis) (July/August): 99–118.

Crowe, C. 2010. "Testing the Transparency Benefits of Inflation Targeting: Evidence from Private Sector Forecasts." *Journal of Monetary Economics* 57 (2): 226–32.

Cukierman, A. 2009. "The Limits of Transparency." *Economic Notes* (Banca Monte dei Paschi di Siena SpA) 38 (1–2): 1–37.

Cukierman, A., and A. H. Meltzer. 1986. "A Theory of Ambiguity, Credibility, and Inflation under Discretion and Asymmetric Information." *Econometrica* 54 (5): 1099–1128.

Dincer, N., and B. Eichengreen. 2010. "Central Bank Transparency: Causes, Consequences and Updates." *Theoretical Inquiries in Law* 11 (75): 75–123.

Dincer, N. N., and B. Eichengreen. 2014. "Central Bank Transparency and Independence: Updates and New Measures." *International Journal of Central Banking* 10 (1, March): 189–253.

Dovern, J., U. Fritsche, and J. Slacalek. 2012. "Disagreement Among Forecasters in G7 Countries." *Review of Economics and Statistics* 94 (4): 1081–96.

Dreher, A., J.-E. Sturm, and J. de Haan. 2010. "When Is a Central Bank Governor Replaced? Evidence Based on a New Data Set." *Journal of Macroeconomics* 32 (3): 766–81.

Ehrmann, M., S. Eijffinger, and M. Fratzscher. 2012. "The Role of Central Bank Transparency for Guiding Private Sector Forecasts." *Scandinavian Journal of Economics* 114 (3): 1018–52.

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

Faust, J., and L. E. O. Svensson. 2001. "Transparency and Credibility: Monetary Policy with Unobservable Goals." *International Economic Review* 42 (2): 369–97.

Freedman, C., and D. Laxton. 2009. "Inflation Targeting Pillars: Transparency and Accountability." IMF Working Paper No. 262.

Geraats, P. M. 2006. "Transparency of Monetary Policy: Theory and Practice." *CESifo Economic Studies* 52 (1): 111–52.

Haldane, A. G. 2017. "A Little More Conversation — A Little Less Action." Speech at the Federal Reserve Bank of San Francisco Macroeconomics and Monetary Policy Conference, March 31.

Hellwig, C., and L. Veldkamp. 2009. "Knowing What Others Know: Coordination Motives in Information Acquisition." *Review of Economic Studies* 76 (1): 223–51.

Jain, M., and C. S. Sutherland. 2018. "How Do Central Bank Projections and Forward Guidance Influence Private-Sector Forecasts?" Staff Working Paper No. 18-2, Bank of Canada.

Jurado, K., S. C. Ludvigson, and S. Ng. 2015. "Measuring Uncertainty." *American Economic Review* 105 (3): 1177–1216.

Kahneman, D. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review* 93 (5): 1449–75.

King, M. 2000. "Monetary Policy: Theory in Practice." Speech at the Meetings of the American Economic Association, Boston, Massachusetts, January 7.

Kuttner, K. N. 2001. “Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market.” *Journal of Monetary Economics* 47 (3): 523–44.

Lange, J., B. Sack, and W. Whitesell. 2003. “Anticipations of Monetary Policy in Financial Markets.” *Journal of Money, Credit and Banking* 35 (6): 889–909.

Lorenzoni, G. 2010. “Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information.” *Review of Economic Studies* 77 (1): 305–38.

Lustenberger, T., and E. Rossi. 2017a. “Does Central Bank Transparency and Communication Affect Financial and Macroeconomic Forecasts?” SNB Working Paper No. 2017-12.

———. 2017b. “The Social Value of Information: A Test of a Beauty and Non-Beauty Contest.” SNB Working Paper No. 2017-17.

Mankiw, G. N., R. Reis, and J. Wolfers. 2003. “Disagreement about Inflation Expectations.” *NBER Macroeconomics Annual 2003,* Vol. 18, ed. M. Gertler and K. Rogoff, 209–48 (chapter 4). MIT Press.

Middeldorp, M. 2011. “Central Bank Transparency, the Accuracy of Professional Forecasts, and Interest Rate Volatility.” Staff Report No. 496, Federal Reserve Bank of New York.

Morris, S., and H. S. Shin. 2002. “Social Value of Public Information.” *American Economic Review* 92 (5): 1521–34.

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 on Survey Forecasts.” *International Journal of Central Banking* 12 (4, December): 147–92.

Neuenkirch, M. 2013. “Central Bank Transparency and Financial Market Expectations: The Case of Emerging Markets.” *Economic Systems* 37 (4): 598–609.

Ranaldo, A., and E. Rossi. 2010. “The Reaction of Asset Markets to Swiss National Bank Communication.” *Journal of International Money and Finance* 29 (3): 486–503.

Santos Silva, J. M. C., and S. Tenreyro. 2006. “The Log of Gravity.” *Review of Economics and Statistics* 88 (4): 641–58.

Sellon, G. H. 2008. “Monetary Policy Transparency and Private Sector Forecasts: Evidence from Survey Data.” *Economic Review* (Federal Reserve Bank of Kansas City) (Third Quarter): 7–34.

Simon, H. A. 1971. “Designing Organizations for an Information-Rich World.” In *Computers, Communications, and the Public Interest*, ed. M. Greenberger, 37–52 (chapter 2). The Johns Hopkins University Press.

Sims, C. A. 2003. “Implications of Rational Inattention.” *Journal of Monetary Economics* 50 (3): 665–90.

Stock, J. H., and M. W. Watson. 2008. “Heteroskedasticity-Robust Standard Errors for Fixed Effects Panel Data Regression.” *Econometrica* 76 (1): 155–74.

Svensson, L. E. O. 2006. “Social Value of Public Information: Comment: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con.” *American Economic Review* 96 (1): 448–52.

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.

Van Der Cruijsen, C. A. B., S. C. W. Eijffinger, and L. H. Hoogduin. 2010. “Optimal Central Bank Transparency.” *Journal of International Money and Finance* 29 (8): 1482–1507.

Wilhelmsen, B.-R., and A. Zaghini. 2011. “Monetary Policy Predictability in the Euro Area: An International Comparison.” *Applied Economics* 43 (20): 2533–44.

Winkler, B. 2002. “Which Kind of Transparency? On the Need for Effective Communication in Monetary Policy Making.” *Ifo Studies* 48 (3): 401–27.

Online Appendix to Does Central Bank Transparency and Communication Affect Financial and Macroeconomic Forecasts?

Thomas Lustenberger[a] and Enzo Rossi[b]
[a]Swiss Financial Market Supervisory Authority (FINMA) and University of Basel
[b]Swiss National Bank and University of Zurich

Appendix B. Tables Referring to Robustness

B.1 Subindexes

Table B.1. Absolute Cross-Sectional Mean Forecast Errors and Subindexes

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)							
	Inflation		**Growth**		**Short-Term Interest Rate**		**Long-Term Interest Rate**	
	CY	**NY**	**CY**	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.019***	0.028***	0.015**	0.013**	0.016***	0.003	0.011**	−0.005
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Transp.								
Political	0.035	0.056	0.084	0.295**	−0.392	−0.430*	0.151	0.097
	(0.13)	(0.14)	(0.12)	(0.12)	(0.26)	(0.25)	(0.17)	(0.25)
Economical	0.096	0.015	0.094	0.148*	−0.129	−0.067	−0.090	−0.201*
	(0.10)	(0.12)	(0.09)	(0.08)	(0.13)	(0.11)	(0.08)	(0.11)
Procedural	0.156	−0.079	−0.021	0.153	0.070	0.062	−0.113	0.148
	(0.11)	(0.11)	(0.09)	(0.11)	(0.21)	(0.21)	(0.09)	(0.16)
Policy	0.006	0.014	−0.011	−0.006	0.118	0.141	0.166**	0.161
	(0.10)	(0.12)	(0.10)	(0.12)	(0.18)	(0.13)	(0.06)	(0.14)
Operational	−0.211	−0.132	−0.004	−0.298**	−0.182	−0.233	−0.163	−0.515*
	(0.14)	(0.14)	(0.17)	(0.12)	(0.22)	(0.20)	(0.13)	(0.26)
Turnover	0.201*	0.189*	0.020	0.028	0.049	0.223**	−0.021	0.002
	(0.11)	(0.10)	(0.07)	(0.10)	(0.13)	(0.10)	(0.08)	(0.11)
VIX	−0.002	0.002	0.001	−0.003**	0.017***	0.002	0.010***	−0.014***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

(*continued*)

Table B.1. (Continued)

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)							
	Inflation		Growth		Short-Term Interest Rate		Long-Term Interest Rate	
	CY	NY	CY	NY	M03	M12	M03	M12
ZLB					−0.124 (0.18)	−0.019 (0.09)	0.080 (0.10)	0.174 (0.11)
N	8,935	8,938	8,679	8,662	4,335	4,344	2,503	2,535
Countries	72	72	70	70	33	33	21	21
R2	0.20	0.18	0.09	0.16	0.15	0.19	0.05	0.17

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of the log of absolute cross-sectional mean forecast errors using subindexes. Inflation and growth have forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. Short-term and long-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year (the intercept, fixed effects, and dummies are not shown in the table).

Table B.2. Forecasts' Cross-Sectional Standard Deviations and Subindexes

	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)							
	Inflation		Growth		Short-Term Interest Rate		Long-Term Interest Rate	
	CY	NY	CY	NY	M03	M12	M03	M12
Speech	0.011***	0.009**	0.002	0.005**	0.009**	0.004	0.007**	0.007***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Political	−0.100	−0.138	0.072	−0.004	−0.394**	−0.240	−0.094	−0.039
	(0.21)	(0.18)	(0.10)	(0.10)	(0.18)	(0.18)	(0.12)	(0.14)
Economical	0.086	0.019	0.046	0.004	0.152	0.139	−0.057	−0.096
	(0.09)	(0.10)	(0.07)	(0.06)	(0.11)	(0.11)	(0.06)	(0.06)
Procedural	−0.222*	−0.191	0.004	−0.045	−0.077	−0.024	−0.116	−0.045
	(0.12)	(0.13)	(0.07)	(0.06)	(0.16)	(0.13)	(0.13)	(0.13)
Policy	0.027	−0.066	0.031	0.035	−0.082	−0.104	0.010	−0.057
	(0.09)	(0.11)	(0.06)	(0.05)	(0.12)	(0.10)	(0.08)	(0.08)
Operational	−0.361**	−0.184	−0.165	−0.099	−0.255	−0.243**	−0.355**	−0.257**
	(0.14)	(0.17)	(0.11)	(0.08)	(0.15)	(0.11)	(0.14)	(0.11)
Turnover	0.109	0.119	0.076	0.042	0.125	0.066	0.101*	0.074
	(0.11)	(0.09)	(0.05)	(0.05)	(0.08)	(0.07)	(0.05)	(0.04)
VIX	0.001	0.006***	0.005***	0.008***	0.015***	0.006***	0.008***	0.004***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ZLB					−0.232	0.059	0.102	0.200
					(0.22)	(0.22)	(0.11)	(0.12)
N	5,383	5,383	5,383	5,383	4,373	4,381	2,786	2,790
Countries	44	44	44	44	33	33	21	21
R2	0.46	0.25	0.50	0.29	0.21	0.25	0.25	0.21

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of the log of cross-sectional standard deviations using subindexes. Inflation and growth have forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. Short-term and long-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year (the intercept, fixed effects, and dummies are not shown in the table).

B.2 Excluding Euro-Area Countries except Germany

Table B.3. Absolute Cross-Sectional Mean Forecast Errors and Forecasts' Cross-Sectional Standard Deviations Excluding Euro-Area Countries Except Germany

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)							
	Inflation		**Growth**		**Short-Term Interest Rate**		**Long-Term Interest Rate**	
	CY	**NY**	**CY**	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.024^{***}	0.029^{***}	0.013^{*}	0.013	0.011	0.001	0.006	0.000
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Transp.	0.049	0.039	0.025	0.074^{*}	−0.010	−0.003	−0.036	-0.094^{*}
	(0.04)	(0.04)	(0.03)	(0.04)	(0.07)	(0.06)	(0.04)	(0.05)
Turnover	0.281^{***}	0.129	0.050	0.090	0.158	0.274^{***}	0.110	0.001
	(0.09)	(0.09)	(0.07)	(0.10)	(0.10)	(0.09)	(0.08)	(0.10)
VIX	−0.004	0.003^{**}	0.002	-0.003^{*}	0.013^{***}	−0.000	0.013^{***}	-0.013^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ZLB					-0.619^{***}	-0.654^{***}	−0.020	0.089
					(0.19)	(0.12)	(0.12)	(0.09)
N	9,760	9,739	9,240	9,163	4,915	4,944	2,701	2,722
Countries	58	58	58	58	29	29	18	18
R2	0.21	0.14	0.12	0.13	0.19	0.20	0.07	0.17

(*continued*)

Table B.3. (Continued)

	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)							
	Inflation		**Growth**		**Short-Term Interest Rate**		**Long-Term Interest Rate**	
	CY	**NY**	**CY**	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.013**	0.011	0.002	0.003	0.019**	0.010**	−0.002	−0.000
	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Transp.	−0.101*	−0.102**	−0.015	−0.020	−0.089	−0.085*	−0.093*	−0.087*
	(0.05)	(0.05)	(0.03)	(0.02)	(0.06)	(0.05)	(0.05)	(0.04)
Turnover.	0.128	0.119	0.071	0.050	0.149	0.068	0.098**	0.062*
	(0.10)	(0.07)	(0.05)	(0.04)	(0.09)	(0.07)	(0.04)	(0.03)
VIX	0.001	0.006***	0.005***	0.008***	0.012***	0.004***	0.009***	0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ZLB					−0.411*	−0.181	0.063	0.136*
					(0.20)	(0.16)	(0.07)	(0.08)
N	6,189	6,189	6,189	6,189	4,981	5,002	2,987	2,991
Countries	36	36	36	36	29	29	18	18
R2	0.45	0.25	0.49	0.30	0.11	0.34	0.20	0.18

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of the log of absolute cross-sectional mean forecast errors and the log of cross-sectional standard deviations excluding euro-area countries except Germany. Inflation and growth have forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. Short-term and long-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year (the intercept, fixed effects, and dummies are not shown in the table).

B.3 Replace the VIX with Jurado, Ludvigson, and Ng (2015)

Table B.4. Absolute Cross-Sectional Mean Forecast Errors and Forecasts' Cross-Sectional Standard Deviations Using Jurado, Ludvigson, and Ng (2015)

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)			
	Inflation	**Growth**	**Short-Term Interest Rate**	**Long-Term Interest Rate**
	CY	**CY**	**M03**	**M03**
Speech	0.021^{***}	0.017^{***}	−0.005	0.018^{***}
	(0.01)	(0.01)	(0.01)	(0.00)
Transp.	0.010	0.021	−0.017	−0.071
	(0.04)	(0.03)	(0.07)	(0.06)
Turnover	0.280^{***}	0.015	0.058	−0.042
	(0.08)	(0.07)	(0.10)	(0.10)
Jurado, Ludvigson, and Ng (2015)	-1.224^{***}	0.510	0.108	-0.555^{**}
	(0.46)	(0.46)	(0.18)	(0.23)
ZLB			-0.655^{***}	0.091
			(0.19)	(0.17)
N	3,169	3,029	1,488	900
Countries	73	73	34	23
R2	0.14	0.08	0.26	0.22

(*continued*)

Table B.4. (Continued)

	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)			
	Inflation	**Growth**	**Short-Term Interest Rate**	**Long-Term Interest Rate**
	CY	**CY**	**M03**	**M03**
Speech	0.014^{***}	−0.006	0.010^{**}	0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Transp.	-0.118^{**}	−0.010	−0.067	-0.142^{**}
	(0.05)	(0.03)	(0.05)	(0.05)
Turnover.	0.128	0.033	0.155^{*}	0.148^{*}
	(0.09)	(0.06)	(0.08)	(0.08)
Jurado, Ludvigson, and Ng (2015)	0.002	−0.071	0.149	-0.157^{***}
	(0.31)	(0.25)	(0.37)	(0.05)
ZLB			−0.238	0.193^{**}
			(0.24)	(0.09)
N	1,914	1,914	1,508	972
Countries	45	45	34	23
R2	0.35	0.35	0.12	0.22

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of the log of absolute cross-sectional mean forecast errors and the log of cross-sectional standard deviations replacing VIX with Jurado, Ludvigson, and Ng (2015). Inflation and growth have forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. Short-term and long-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year (the intercept, fixed effects, and dummies are not shown in the table).

B.4 Benchmark Model with Only Basic Central Bank

Table B.5. Absolute Cross-Sectional Mean Forecast Errors and Forecasts' Cross-Sectional Standard Deviations Using Basic Central Bank Sample

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)							
	Inflation		Growth		Short-Term Interest Rate		Long-Term Interest Rate	
	CY	**NY**	**CY**	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.015**	0.024***	0.011*	0.006	0.010**	0.002	0.006	0.001
	(0.01)	(0.01)	(0.01)	(001)	(0.00)	(0.00)	(0.00)	(0.00)
Transp.	−0.034	−0.041	0.064	0.018	0.006	0.024	−0.043	−0.106*
	(0.08)	(0.11)	(0.07)	(0.07)	(0.09)	(0.09)	(0.03)	(0.05)
Turnover	0.267***	0.090	0.167*	0.143	0.050	0.178	0.031	0.010
	(0.08)	(0.12)	(0.08)	(0.13)	(0.10)	(0.10)	(0.10)	(0.10)
VIX	0.002	−0.001	0.002	−0.003	0.017***	−0.002	0.010***	−0.009***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ZLB					−0.737***	−0.669***	0.110	0.220
					(0.17)	(0.13)	(0.12)	(0.13)
N	5,458	5,458	5,398	5,362	4,376	4,398	3,043	3,062
Countries	28	28	28	28	23	23	17	17
R2	0.25	0.18	0.14	0.19	0.24	0.26	0.08	0.21

(*continued*)

Table B.5. (Continued)

	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)							
	Inflation		**Growth**		**Short-Term Interest Rate**		**Long-Term Interest Rate**	
	CY	**NY**	**CY**	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.015***	0.009*	0.003	0.008**	0.016***	0.002	0.008**	0.005**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Transp.	−0.089	−0.084	−0.021	−0.012	−0.066	−0.057	−0.116	−0.105
	(0.10)	(0.09)	(0.05)	(0.04)	(0.08)	(0.08)	(0.08)	(0.06)
Turnover.	0.055	0.061	0.016	0.047	0.155*	0.037	0.138**	0.062
	(0.06)	(0.05)	(0.04)	(0.04)	(0.07)	(0.06)	(0.05)	(0.05)
VIX	0.001	0.005***	0.004***	0.007***	0.013***	0.005***	0.010***	0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ZLB					−0.271	−0.162	0.109	0.155*
					(0.21)	(0.15)	(0.08)	(0.08)
N	4,636	4,636	4,636	4,636	4,438	4,456	3,231	3,231
Countries	24	24	24	24	23	23	17	17
R2	0.49	0.20	0.52	0.27	0.11	0.34	0.21	0.18

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of the log of absolute cross-sectional mean forecast errors and the log of cross-sectional standard deviations using the basic central bank sample. Inflation and growth have forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. Short-term and long-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year (the intercept, fixed effects, and dummies are not shown in the table).

B.5 Benchmark Model Excluding Months with Zero Speeches

Table B.6. Absolute Cross-Sectional Mean Forecast Errors and Forecasts' Cross-Sectional Standard Deviations Excluding Months with Zero Speeches

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)							
	Inflation		**Growth**		**Short-Term Interest Rate**		**Long-Term Interest Rate**	
	CY	**NY**	**CY**	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.010**	0.015***	0.007	0.005	0.005	0.001	0.003	0.002
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Transp.	−0.080	−0.032	0.103**	−0.028	0.011	0.015	−0.078**	−0.074
	(0.05)	(0.08)	(0.05)	(0.06)	(0.08)	(0.08)	(0.03)	(0.05)
Turnover	0.149**	0.008	−0.124	−0.065	−0.100	0.062	0.024	0.012
	(0.07)	(0.11)	(0.10)	(0.13)	(0.09)	(0.09)	(0.13)	(0.13)
VIX	0.001	−0.003	0.004	−0.003	0.018***	−0.001	0.011***	−0.015***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ZLB					−0.636***	−0.681***	0.146	0.244*
					(0.17)	(0.13)	(0.11)	(0.13)
N	5,352	5,355	5,311	5,296	3,335	3,338	2,657	2,671
Countries	58	58	57	57	32	32	22	22
R2	0.24	0.21	0.12	0.21	0.25	0.29	0.08	0.21

(*continued*)

Table B.6. (Continued)

	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)							
	Inflation		Growth		Short-Term Interest Rate		Long-Term Interest Rate	
	CY	NY	CY	NY	M03	M12	M03	M12
Speech	0.009***	0.003	0.000	0.004*	0.007**	−0.001	0.007**	0.004**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Transp.	−0.096	−0.086	−0.009	−0.019	−0.057	−0.045	−0.059	−0.054
	(0.07)	(0.07)	(0.02)	(0.03)	(0.08)	(0.07)	(0.04)	(0.03)
Turnover.	−0.002	−0.005	−0.020	−0.015	0.021	−0.055	0.081	0.003
	(0.04)	(0.05)	(0.03)	(0.03)	(0.06)	(0.05)	(0.06)	(0.05)
VIX	0.003**	0.006***	0.004***	0.008***	0.018***	0.007***	0.010***	0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ZLB					−0.507***	−0.283**	0.113	0.129*
					(0.14)	(0.11)	(0.08)	(0.07)
N	3,594	3,594	3,594	3,594	3,371	3,374	2,737	2,739
Countries	40	40	40	40	32	32	22	22
R2	0.56	0.17	0.35	0.25	0.19	0.32	0.16	0.11

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of the log of absolute cross-sectional mean forecast errors and the log of cross-sectional standard deviations excluding months with zero central bank speeches. Inflation and growth have forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. Short-term and long-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year (the intercept, fixed effects, and dummies are not shown in the table).

B.6 Benchmark Model Correcting for Season (excluding July, August, and December)

Table B.7. Absolute Cross-Sectional Mean Forecast Errors and Forecasts' Cross-Sectional Standard Deviations Correcting for Season

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)							
	Inflation		**Growth**		**Short-Term Interest Rate**		**Long-Term Interest Rate**	
	CY	**NY**	**CY**	**NY**	**M03**	**M12**	**M03**	**M12**
Speech	0.019***	0.021***	0.014***	0.010	0.011**	0.006	0.009*	0.005
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Transp.	0.016	−0.009	0.036	0.050	−0.054	−0.018	−0.061	−0.095*
	(0.03)	(0.04)	(0.03)	(0.04)	(0.06)	(0.06)	(0.04)	(0.05)
Turnover	0.257***	0.071	0.057	0.074	0.097	0.288***	0.015	0.032
	(0.07)	(0.08)	(0.07)	(0.09)	(0.12)	(0.08)	(0.09)	(0.10)
VIX	−0.002	0.003*	0.002	−0.003*	0.018***	0.001	0.017***	−0.010**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ZLB					−0.687***	−0.762***	0.075	0.298**
					(0.18)	(0.14)	(0.11)	(0.12)
N	9,277	9,260	8,882	8,825	4,378	4,401	2,670	2,688
Countries	73	73	73	73	34	34	23	23
R2	0.21	0.15	0.12	0.14	0.20	0.23	0.09	0.20

(*continued*)

Table B.7. (Continued)

	Cross-Sectional Standard Deviations of Forecasts (Cross-Sectional Standard Deviations of Forecasts at Each Point in Time)							
	Inflation		Growth		Short-Term Interest Rate		Long-Term Interest Rate	
	CY	NY	CY	NY	M03	M12	M03	M12
Speech	0.013^{***}	0.008^{**}	−0.000	0.004	0.015^{***}	0.004	0.008^{**}	0.005^{*}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Transp.	-0.113^{**}	-0.117^{***}	−0.014	−0.024	-0.091^{*}	-0.084^{**}	-0.101^{*}	-0.086^{**}
	(0.04)	(0.04)	(0.02)	(0.02)	(0.05)	(0.04)	(0.05)	(1.04)
Turnover.	0.123	0.083	0.060	0.054	0.174^{**}	0.077	0.091^{**}	0.053^{*}
	(0.09)	(0.06)	(0.04)	(0.04)	(0.08)	(0.07)	(0.04)	(0.03)
VIX	0.001	0.005^{***}	0.004^{***}	0.007^{***}	0.012^{***}	0.005^{***}	0.009^{***}	0.005^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ZLB					−0.319	−0.187	0.110	0.157^{**}
					(0.19)	(0.14)	(0.008)	(0.07)
N	5,608	5,608	5,608	5,608	4,425	4,444	2,887	2,891
Countries	45	45	45	45	34	34	23	23
R2	0.43	0.25	0.48	0.28	0.11	0.34	0.21	0.16

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of the log of absolute cross-sectional mean forecast errors and the log of cross-sectional standard deviations excluding July, August, and December. Inflation and growth have forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. Short-term and long-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year (the intercept, fixed effects, and dummies are not shown in the table).

B.7 Benchmark Model but Excluding Outliers

Table B.8. Absolute Cross-Sectional Mean Forecast Errors but Excluding Outliers

	Absolute Cross-Sectional Mean Forecast Errors (Absolute Differences between the Cross-Sectional Mean Forecasts and Their Realizations)			
	Inflation		**Short-Term Interest Rate**	
	CY	**NY**	**M03**	**M12**
Speech	0.016***	0.018***	0.006	0.000
	(0.00)	(0.01)	(0.00)	(0.00)
Transp.	0.010	−0.024	−0.016	0.007
	(0.03)	(0.05)	(0.06)	(0.06)
Turnover	0.176***	−0.022	0.046	0.266***
	(0.06)	(0.08)	(0.09)	(0.09)
VIX	−0.002	0.002	0.017***	−0.001
	(0.00)	(0.00)	(0.00)	(0.00)
ZLB			−0.680***	−0.714***
			(0.16)	(0.13)
N	11,138	11,129	5,354	5,384
Countries	65	65	31	31
R2	0.21	0.15	0.22	0.24

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of the log of absolute cross-sectional mean forecast errors excluding outliers. For inflation we excluded ARG, BGR, BLR, MDA, RUS, TUR, UKR, and VEN. For short-term interest rates we excluded ARG, IDN, and VEN. Inflation has forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. Short-term interest rates have two forecast horizons. M03 is the 3-month forecast horizon, while M12 stands for the 12-month forecast horizon. We also include a dummy for each year (the intercept, fixed effects, and dummies are not shown in the table).

B.8 Replace Realized GDP Figure with December Forecast of the Current Year

Table B.9. Absolute Cross-Sectional Mean Forecast Errors and Realized Real GPD

	Growth	
	CY	**NY**
Speech	0.003	0.003
	(0.01)	(0.01)
Transp.	−0.072*	0.020
	(0.04)	(0.04)
Turnover	0.023	0.019
	(0.06)	(0.09)
VIX	0.004**	−0.002**
	(0.00)	(0.00)
N	11,104	12,266
Countries	73	73
R2	0.25	0.19

Notes: Country fixed-effects panel regression with panel clustered standard errors in parentheses. *, **, and *** denote $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The table shows results for a fixed-effects panel regression of the log of absolute cross-sectional mean forecast errors using the December forecast of the current year as realized growth figure. Growth has forecasts for the current-year (CY) and the next-year (NY) value. We include a dummy for each forecast horizon since both current-year and next-year forecasts' horizons decrease over time. We also include a dummy for each year (the intercept, fixed effects, and dummies are not shown in the table).

Stress Tests of the Household Sector Using Microdata from Survey and Administrative Sources*

Jaanika Meriküll and Tairi Rõõm
Eesti Pank

This paper conducts microsimulation-based stress tests to assess the financial risks of the household sector. The Estonian Household Finance and Consumption Survey data set is employed, where the survey data from household interviews are complemented with the same information from administrative registers. We analyze the sensitivity of financial-sector loan losses to adverse shocks. It is found that the survey data and the register data indicate the same segment of vulnerable households. The main difference between the two data sources is that the losses predicted by the register data are larger. This is mostly the result of the overestimation of assets and underestimation of liabilities in the survey.

JEL Codes: D14, E43, G21, C81.

1. Introduction and Related Literature

Household borrowing has increased considerably in most European countries in recent decades, both in absolute terms and in relation to household income. The rapid increase in household debt in the United States was one of the triggers of the Great Recession (e.g., Mian and Sufi 2010). These developments have been the source of concerns about the financial system and have led many central banks to undertake stress-testing exercises to assess the resilience of the household sector to various macroeconomic

*The authors would like to thank Hanno Kase, Jana Kask, Merike Kukk, and Olavi Miller for their insightful comments on the earlier drafts of this article. The views expressed are those of the authors and do not necessarily represent the official views of Eesti Pank or the Eurosystem. Authors' e-mail addresses: jaanika.merikyll@eestipank.ee, tairi.room@eestipank.ee.

risks. Earlier analyses of financial stability were almost exclusively based on macrodata, but recently they have increasingly been conducted on disaggregated data. The use of microlevel statistics on household finances has several advantages over macrolevel studies. It allows the extent of households' financial buffers to be assessed together with their liabilities and makes it possible to evaluate the distribution of the debt burden across different household types.

As part of the initiative to base financial stability analysis on disaggregated data, the euro-area central banks together with the European Central Bank launched the Household Finance and Consumption Survey. This survey collects data on households' assets and liabilities in a harmonized manner across all the euro-area countries. These data are representative at both the national and euro-area levels and contain comprehensive information on households' balance sheets. The availability of the Household Finance and Consumption Survey data makes it possible to conduct microdata-based studies of households' ability to service their debts and of the credit risks to the financial sector that emanate from the household sector. Several recent studies focus on financial fragility analysis using the Household Finance and Consumption Survey data or analogous microlevel data sets.[1]

While most of the countries collect household balance sheet data by interviewing households, there is an increasing tendency to compile such databases on the basis of administrative sources.[2] The aim

[1]Examples of such studies include Johansson and Persson (2006) for Sweden; Herrala and Kauko (2007) for Finland; Holló and Papp (2007) for Hungary; Albacete and Fessler (2010) for Austria; Faruqui, Liu, and Roberts (2012) for Canada; Martinez et al. (2013) for Chile; Michelangeli and Pietrunti (2014) for Italy; Bańbula et al. (2015) for Poland; Bilston, Johnson, and Read (2015) for Australia; Ampudia, van Vlokhoven, and Zochowski (2016) for 10 euro-area countries; and Galuščák, Hlaváč, and Jakubík (2016) for the Czech Republic.

[2]Finland is the only country that relies mostly on administrative register data in the Household Finance and Consumption Survey (Eurosystem Household Finance and Consumption Network 2016). Register data-based household stress testing is already used by the Danish central bank (Andersen et al. 2013). The European Central Bank plans to launch an analytical credit data set (AnaCredit) that will collect whole population-based corporate- and housing-sector loan-level data that can be used for stress testing (Israel et al. 2017).

of our paper is to construct households' financial fragility indicators and conduct stress-testing exercises using and comparing data from different sources, i.e., from interviews and from administrative registers. To the best of our knowledge, ours is the first paper to conduct such a comparison. This lets us evaluate whether different data-collection methods yield different messages in the assessment of the financial risks of the household sector.

As a general rule, the Household Finance and Consumption Survey data were collected by household interviews, but some euro-area countries partially complemented or replaced survey data with variables from administrative sources. While this was mostly done only partially, the Household Finance and Consumption Survey data set for one country, Estonia, contains comprehensive information on households' assets, liabilities, and incomes from two separate sources, i.e., from administrative files as well as from the survey. We employ the Estonian Household Finance and Consumption Survey data set to conduct stress tests of households' financial fragility using both the survey data and the administrative records.

Being able to access the data from these two alternative sources lets us evaluate possible biases in the survey-based estimates of household incomes, liabilities, and assets relative to the administrative data. While there are many studies which compare incomes from surveys and administrative sources (e.g., Duncan and Hill 1985; Pischke 1995; Bound, Brown, and Mathiowetz 2001; Kapteyn and Ypma 2007; Chen, Hong, and Nekipelov 2011), there are only a few papers that provide a similar comparison for households' net wealth (e.g., Johansson and Klevmarken 2007). A comparison of liabilities is also given in only a limited number of studies (e.g., Brown et al. 2011). To the best of our knowledge, there is no study that uses all the information about income, assets, and liabilities and compiles financial fragility indicators for households from survey and register data.

This paper conducts stress tests of the household sector by quantifying the effect of various adverse shocks on household finances and by assessing the ability of households to continue servicing their debts after they have been exposed to shocks. In addition, we look at the distributional aspects of financial distress by detecting groups of households which are more financially fragile or are particularly vulnerable to adverse shocks. The estimation methodology that we use in the stress-testing exercises builds on approaches that were used by the earlier studies, which employed microlevel data to assess the financial risks of the household sector.

The earlier papers that evaluate households' financial vulnerability can be divided into three groups. The first approach assesses financial fragility by finding the fraction of households whose ratio of debt service to income exceeds some threshold level. Most typically, households are considered financially fragile when their ratio of debt service to income exceeds 30 percent, but some authors also use 35 percent or 40 percent. Examples of the studies using this approach include Faruqui, Liu, and Roberts (2012), Martinez et al. (2013), and Michelangeli and Pietrunti (2014).

The second method is based on the share of households with a negative financial margin. The financial margin is defined as current disposable income from which essential consumption expenditures and loan servicing costs are deducted. In these studies, it is typically assumed that all households with negative financial margins will default on their loans. This approach is used in, e.g., Johansson and Persson (2006), Holló and Papp (2007), and Albacete and Fessler (2010).

The third method considers the share of liquid assets the household has in addition to the financial margin when determining the likelihood of the household defaulting on its loans. This method is used, e.g., in recent studies by Ampudia, van Vlokhoven, and Zochowski (2014, 2016). Earlier studies of household stress tests assumed that the probability of defaulting on loans equals one for all households with a negative financial margin. However, in practice, only some households with a current negative financial margin default on their loans, since the probability of default is also dependent on the extent of the financial buffers they have in the form of liquid assets. In the approach used by Ampudia, van Vlokhoven, and Zochowski (2014, 2016), households will default on their loans only when they have negative financial margins and when the amount of liquid assets they have is smaller than a particular threshold value. This threshold value is calibrated so that the aggregate share of defaulted loans is equal to the actual share of nonperforming household loans in the banking sector for the same time period.

Our stress-testing methodology is closest to that used in the study by Ampudia, van Vlokhoven, and Zochowski (2016).

Specifically, we follow their idea of calibrating the probabilities of default for households so that the microdata-based exposure at default matches the aggregate historical share of households' non-performing loans (NPLs) in the banking sector.

The structure of the stress-testing exercises that is employed in this paper is similar to that in earlier studies which conducted stress tests. We use the following steps. First, the probability of default on household loans is assessed using survey and register data. Second, we simulate the effects of adverse shocks of one, two, and three standard deviations. The variables to which the shocks are applied are the base interest rate, the unemployment rate, and real estate prices. The next step is to assess the effect of these shocks on households' probability of default (PD). Finally, the resulting effect on the banks' household loan portfolios is evaluated by calculating the effect on the share of loans exposed to default (the exposure at default, or EAD) and on the resulting losses from defaulted loans (the loss given default, or LGD). The stress-testing exercises are performed using the survey data as well as the data from administrative files, which lets us compare the sensitivity of the results to the use of different data sources.

The biases in the survey data of the Estonian Household Finance and Consumption Survey are similar in many ways to those detected by the earlier studies, but we also identify some differences. First, we evaluate the level of indebtedness on the basis of the survey and compare it with data from the administrative files. We find no systematic bias in the reporting of the outstanding amounts of mortgage loans, i.e., survey and administrative records yield similar estimated outstanding amounts of mortgage loans on average. A similar comparison for noncollateralized loans shows that their amounts are underestimated in the survey relative to administrative records. This finding is in correspondence with the results of an earlier study in which data from the U.S. Survey of Consumer Finances are compared with administrative records, and which also shows that consumer loans tend to be underreported in surveys (Brown et al. 2011).

The estimated value of household assets is mainly determined by their real estate holdings, as real estate is by far the largest component of the assets of the average household (Eurosystem Household

Finance and Consumption Network 2013b). Comparing the register-based estimates with those reported by the survey respondents shows that households tend to overestimate the value of their real estate in the Estonian Household Finance and Consumption Survey. This result is in line with findings in some previous studies, which also imply that real estate values are upward biased in household surveys (e.g., Kiel and Zabel 1999, Johansson and Klevmarken 2007). Interestingly, conducting these comparisons for different types of real estate in the Estonian Household Finance and Consumption Survey implies that households tend to overestimate the value of their main residence, whereas their other real estate holdings are undervalued in the survey.[3]

In addition, we are able to compare the survey-based estimates of household incomes in the Estonian Household Finance and Consumption Survey with estimates based on the register data. Earlier studies have mostly found that incomes tend to be underreported in surveys, but the measurement error is usually negatively correlated with income level, i.e., low-income households tend to overreport their incomes and high-income households tend to underreport theirs in the survey (e.g., Rodgers, Brown, and Duncan 1993; Bound et al. 1994; and Pischke 1995). The negative correlation between the measurement error and the level of income is also present in the Estonian Household Finance and Consumption Survey. In the Estonian Household Finance and Consumption Survey, however, the register-based estimates of income are, on average, smaller than the survey-based estimates, which is an atypical finding.

Since survey-based estimates of income and assets are overvalued while liabilities are undervalued in the Estonian Household Finance and Consumption Survey, replacing the survey data with register data results in larger estimated measures of financial distress for the household sector and larger losses from defaults on household debts for the banking sector. The estimated losses for banks are up to four times larger in the stress tests based on register data than according to the survey data. Both data sources point to the same

[3]Main residences make up by far the largest part of households' real estate holdings. Therefore the overvaluation effects dominate and, on average, real estate holdings are overvalued.

segment of vulnerable households, indicating that low-income households bear most of the risks from household debts. The probability of default and the vulnerability to shocks is more concentrated to low-income households according to the survey data than it is according to the register data. These findings have potential implications for other similar studies. As a general rule, most of the earlier analyses of financial fragility have been based on survey data, and so it is likely that these evaluations underestimated the financial risks of the household sector. It is also relevant for cross-country comparisons, since if some of the country studies are based on administrative files and others on survey data, then the results are not directly comparable.

The paper is structured as follows. The second section describes the Estonian Household Finance and Consumption Survey data used for the analysis. The third section presents the derivation of the measures of financial fragility for households that are used in the stress tests and a comparison of these variables based on survey and administrative data. The fourth section presents the results of the stress-testing exercises using the survey and administrative data, and the last section provides the conclusions.

2. The Data

The Household Finance and Consumption Survey data set contains detailed household-level data on various items of household balance sheets together with related demographic and economic variables, including various types of income, employment status, inheritances and gifts, consumption, etc. The unique feature of the Estonian Household Finance and Consumption Survey data is that the information collected by the survey was complemented by information from administrative registers, which allows the data from different sources to be compared and potential biases in reporting to the surveys to be estimated. There is no reason to believe that the results based on the survey and administrative data are unique to Estonia. Estonia is a representative case because the collection of the survey data was well harmonized with the other euro-area countries and was undertaken by the National Statistical Institution. Estonia is a European Union (EU) member state with a harmonized institutional setting of financial intermediation, and the level of its debt burden is at the average among the euro-area countries, although household debts have been accumulated more recently than in most of the other euro-area countries (Meriküll and Rõõm 2016, Kukk 2019).

The fieldwork for the Estonian Household Finance and Consumption Survey took place between March and June 2013, and the final sample contains 2,220 households. The sampling design was one-stage stratified systematic sampling. Wealthy households were oversampled in the survey to give better coverage of households' assets. The survey mode was computer-assisted personal interviews. The estimation weights were calculated to adjust for survey nonresponse and were calibrated for age, sex, degree of urbanization, ethnicity, education, household size, and homeownership status. Replicate weights were introduced for variance estimation, and bootstrap methods with replacement were used to create 1,000 replication weights. Multiple stochastic imputation was used to fill in the data for missing observations. The imputation was not applied to the whole survey, but the key variables, such as the components of net wealth, income, and consumption, were imputed. The methodology for calculating the weights and for the imputation was similar to that used in other euro-area countries participating in the Household Finance and Consumption Survey; see Eurosystem Household Finance and Consumption Network (2013a) for more details. A more detailed explanation of the sample statistics of the Estonian Household Finance and Consumption Survey is given in Meriküll and Rõõm (2016).

A large number of survey items were also collected from administrative sources in Estonia. This served more than one purpose. The first aim was to validate the data collected by interviews. Second, the administrative data were used for the imputation of the missing observations. Third, the data from alternative sources were collected with the longer-term objective of gradually replacing the items collected from the interviews with administrative information in future waves of the survey. It is important for the context of this paper that in the editing process of the survey data set, the survey variables were mainly compiled using the information collected by interviews, so that the survey data and administrative data can be meaningfully compared. The only wealth component that was replaced in the survey with administrative data was the amount of financial assets held domestically. Financial assets mainly consist of bank deposits in Estonia and make up 10 percent on average of the gross value of all financial plus real assets.

The main administrative sources of data on disposable income were the Income and Social Security Tax Register, the Social Transfers Register, the Health Insurance Benefits Register, and the Unemployment Fund Register. Data from the Land Board were used for the estimates of real estate prices, data from the commercial banks were used for the outstanding amounts of deposits and financial liabilities, and data from the Central Register of Securities were used for the value of other financial assets. The administrative data were collected by Statistics Estonia and the Bank of Estonia and merged with the interview data collected from the household survey. All the variables from administrative sources were collected at the individual or household level.

The data that we use for comparative purposes are similar to those used in other studies comparing income and wealth data from survey and administrative sources (e.g., Johansson and Klevmarken 2007). The estimated values of financial assets and liabilities that are based on the administrative data are not completely free of measurement error, but they can be expected to be much closer to the true values of these variables than the survey-based estimates. The register-based estimates of income, on the other hand, are more likely to contain systematic measurement errors. The income data from the administrative sources may be underestimated because of tax evasion, as the share of wages that are undeclared has been estimated at 5–20 percent in Estonia (Putnins and Sauka 2015), and the share of undeclared self-employment income has even been put as high as 60 percent (Kukk and Staehr 2014). Furthermore, it is not clear whether respondents report the true value of their income in surveys or give the same income as that reported to the tax office. Households' choices about how much of their true income to report depend partly on survey design (the framing effects, etc.).

The assessment of the value of real estate from the register data is based on the average price of market transactions for different real estate types within a detailed district, which is a similar approach to that taken by the study of Johansson and Klevmarken (2007) using Swedish data.[4] The U.S. practice has been to compare survey-based wealth data that are calculated using the balance sheet principle with a register counterpart that has been indirectly calculated from income tax data using the method of capitalizing income flows (Bricker et al. 2016). However, the focus of similar analysis in the United States has been on comparing the estimate of the share of wealth held by the top 1 percent in various sources and not strictly on comparing the survey and administrative values of the same individual for all the people sampled.

3. Methodology and Estimates of Financial Fragility

3.1 Estimation Methodology

In this subsection we derive the measures of household financial fragility that are used in the stress tests. First, we define the household financial margin (FM). Following this, we show how the household probability of default is calculated on the basis of the FM and liquid assets. The probability of default is calibrated to match the aggregate household-sector ratio of nonperforming loans. Finally, a measure of banking-sector losses (loss given default, LGD) is defined, which provides an estimate of the effect of household-sector loan quality on financial stability.

The household financial margin is derived as follows:

$$FM_i = Y_i - DP_i - C_{i,} \tag{1}$$

where FM_i denotes the financial margin of household i, Y_i is total disposable income, DP_i is total debt service costs, and C_i is essential consumption. Total disposable income covers the after-tax income of all household members from all sources, covering labor income, capital income, pensions, and any other public or private transfers. Income is collected for the calendar year preceding the survey (2012) and is divided by 12 to obtain average monthly income. The data are collected in gross terms and converted to net terms using statutory

[4]They employ a slightly more complicated approach based on the average transaction price in a region and the tax-assessed value.

tax rates and exemptions.[5] Debt payments DP_i consist of monthly payments for mortgages and other loans; other loans are all consumer loans and loans from employers or other households, except leases, credit line overdrafts, and credit card debt.[6] The reference period is the time of the survey, and payments cover only interest and loan principal payments, but do not cover insurance, taxes, or other fees.

Essential consumption or basic consumption C_i has been defined as the Statistics Estonia official estimate of the subsistence minimum (Statistics Estonia, table hh27 at http://www.stat.ee). The subsistence minimum without expenditures on housing was EUR 128 for single-person households in 2013. The subsistence minimum for households with more than one member is calculated by multiplying this amount by the sum of consumption weights taken from the OECD (Organisation for Economic Co-operation and Development) equivalence scale.[7] We add the monthly rental payments to the subsistence minimum to calculate the total level of basic consumption for renters.

Authors of earlier studies have taken various approaches to defining essential consumption, with some defining it as the subsistence minimum or poverty line (Bilston, Johnson, and Read 2015;

[5]Although the Estonian tax system is relatively simple, with a flat tax rate and only a few tax exemptions, several assumptions are still required for disposable income to be derived from all the income types at the household level. It is assumed that the tax-exempt amount for total income and the additional exemption for the retired apply, and various deductions have been assumed, including exemptions for household main residence mortgage interest payments, children, and investments in voluntary pension schemes. It is also assumed that no income taxes are paid on rental income or on self-employment income from abroad, as tax evasion is common for these income types. Married couples are assumed to submit joint income declarations. The household member with the highest income is assumed to declare the household-level income and to deduct all the household-level deductibles in households with no married couple.

[6]Leases, credit card debt, and bank account overdrafts are excluded because the data on monthly payments for these loans are not available in the Household Finance and Consumption Survey. The exclusion of these loans should not have a major effect on the results since the majority of the loan burden in Estonia consists of mortgages, and collateralized loans make up 95 percent of the total loan burden excluding leases.

[7]The first adult household member gets a weight of 1, each subsequent household member who is at least 14 years old gets a weight of 0.5, and each household member aged less than 14 gets a weight of 0.3.

Ampudia, van Vlokhoven, and Zochowski 2016) or as the household self-reported minimum subsistence level (Albacete and Fessler 2010), and some defining it more generously as consumption of food, energy, health, and rent (Galuščák, Hlaváč, and Jakubík 2016) or the minimum nondurable consumption and non-interest housing costs (Johansson and Persson 2006). We prefer to use the subsistence minimum instead of the actual expenditures on the most essential consumption categories because it is likely that consumption is reduced in response to negative shocks.

Most studies of household stress tests consider all households with a negative financial margin as distressed households and define their probability of default as equal to one. However, in practice, only some households with a current negative financial margin default on loans, since the probability of default is also dependent on financial buffers. Households with a substantial level of liquid assets may be able to cover the negative financial margin for some time until they manage to restore their income and so can avoid default. This paper applies the solvency and liquidity approach introduced by Ampudia, van Vlokhoven, and Zochowski (2016) to derive the probability of default. They show that this type of distress measure outperforms other indicators that are based on a negative financial margin or on debt-service-to-income ratio thresholds, as these tend to overestimate the exposure at default. The method used by Ampudia, van Vlokhoven, and Zochowski (2016) not only has a more realistic distress measure that employs information on income as well as on liquid assets, but also allows flexible calibration of the exposure at default ratio so that it meets the actual aggregate nonperforming loan ratio. As a result, stress tests based on microdata and macrodata can easily be compared at the same meaningful scale.

Using the idea from Ampudia, van Vlokhoven, and Zochowski (2016), we define the probability of default as follows:

$$\begin{gathered} \textit{If } FM_i \geq 0 \textit{ then } pd_i = 0 \\ \textit{If } FM_i < 0 \wedge LIQ_i \geq |FM_i| \times M \textit{ then } pd_i = 0 \\ \textit{If } FM_i < 0 \wedge 0 < LIQ_i < |FM_i| \times M \textit{ then } pd_i = 1 - \tfrac{LIQ_i}{|FM_i|} \times \tfrac{1}{M} \\ \textit{If } FM_i < 0 \wedge LIQ_i = 0 \textit{ then } pd_i = 1, \end{gathered} \tag{2}$$

where pd_i denotes the probability of default of household i, FM_i is the financial margin, LIQ_i are liquid assets, and M is the calibrated number of months after which the household restores its non-negative financial margin. Equation (2) assumes that M is greater than zero. Liquid assets are household net liquid assets, i.e., the sum of deposits, mutual funds, bonds, non-self-employment business wealth, publicly traded shares, and managed accounts from which bank overdraft debts and credit card debts are deducted.[8] The very first line of the set of equations (2) shows that households with a positive financial margin will not default and their probability of default is zero. Households with a negative financial margin and enough liquid assets to cover the calibrated M months of the negative financial margin will also not default. Households with a negative financial margin and no liquid assets will default with the probability of one, while households in between these two extremes have a probability of default that is a decreasing linear function of the ratio of liquid assets to the absolute value of the financial margin.

After obtaining the estimated probabilities of default for the households, we calculate the banks' exposure at default or the share of defaulting loans in the total loan stock. Following Ampudia, van Vlokhoven, and Zochowski (2016), the formula for calculating EAD is

$$EAD = \frac{\sum_{i=1}^{N} pd_i D_i}{\sum_{i=1}^{N} D_i}, \tag{3}$$

where EAD denotes exposure at default and D_i is the total debt of household i. The value of M is calibrated so that the estimated EAD would meet the aggregate share of nonperforming loans in Estonia at the time of the survey, i.e., from March to June 2013. The survey data are used for calibration of M, and the same value of M is used for estimations based on register data. The aggregate NPL share was assessed as the percentage of household loans in the total loan stock whose payments were past due for more than 30 days, which was 3.4 percent during the survey fieldwork period (Bank of Estonia statistics table 3.3.11). Ampudia, van Vlokhoven,

[8] These credit types are not taken into account in calculations of the financial margin.

and Zochowski (2016) calibrate the value of M to meet the nonperforming loan ratio for the euro-area households and find M to vary a lot across countries, from 0 to 26 months.[9] Given that the share of households with a negative financial margin in the Estonian data is high compared with the actual NPL ratio (13.0 percent versus 3.4 percent; see table 1), the value of M has to be relatively low in Estonia, meaning that although the negative financial margin is frequently encountered, households can restore their financial solvency relatively quickly. The calibration shows that the calibrated value of M is one month, which results in an aggregate value for EAD of 3.4 percent.

Lastly, the share of banks' loan losses that are caused by defaults, or the loss given default, can be calculated as the probability of default multiplied by the sum of potential loan losses for mortgage loans with negative equity and the sum of all noncollateralized loans (following the idea of Herrala and Kauko 2007 and the notation of Ampudia, van Vlokhoven, and Zacchowsky 2016):

$$LGD = \frac{\sum_{i=1}^{N} pd_i[(D_i^M - W_i^M)c_i^M + D_i^{NC}]}{\sum_{i=1}^{N} D_i}, \tag{4}$$

where LGD denotes loss given default, Di denotes debt, superscript M denotes mortgage loans, superscript NC denotes noncollateralized loans, W_i denotes assets that the bank can liquidate in the event of a default, and c_i equals one if the household is "underwater," meaning its collateral has a lower value than the outstanding value of its loans, while c_i is zero otherwise. The value of W_i is taken as the value of all the real estate assets that a given household owns. The LGD provides an estimate of the potential losses for banks from nonperforming loans.

3.2 Households' Financial Fragility Indicators: Survey Data and Historical Aggregates

Table 1 reports descriptive statistics for the financial fragility indicators from the survey: the share of households with a negative

[9]Their study covers the households of the 15 euro-area countries that participated in the first wave of the Household Finance and Consumption Survey, these being all the euro-area member states in 2010 except Ireland.

Table 1. Indicators of the Financial Fragility of Households and Estimated Loan Losses for Banks: Survey Data and Historical Aggregates

	Aggregate Historical Measures[a]	Survey-Based Estimates
Negative Financial Margin, %	—	13.0
Probability of Default, %	—	5.2
Exposure at Default, %	3.4	3.4
Mortgages, %	2.8	3.2
Noncollateralized Loans, %	6.4	8.8
Exposure at Default, Mln. EUR	230.7	165.9
Mortages, Mln. EUR	163.0	147.4
Noncollateralized Loans, Mln. EUR	67.2	18.5
Loss Given Default, %	0.8	0.4
Mortgages, %	0.5	0.0
Noncollateralized Loans, %	2.5	8.8
Loss Given Default, Mln. EUR	55.5	20.6
Mortgages, Mln. EUR	29.8	2.1
Noncollateralized Loans, Mln. EUR	25.8	18.5
No. of Obs.		769

Sources: Authors' calculations from Estonian Household Finance and Consumption Survey data; the Bank Estonia statistics table 3.3.11 for the aggregate nonperforming loans; and the Bank of Estonia credit risk model for loan loss provisions.
Note: Indebted households are defined as households with collateralized debt and with consumer loans, not including leases, credit line overdrafts, and credit card debt.
[a]Exposure at default is measured as the aggregate ratio of nonperforming loans with debt payments 30 days or more past due in the survey period. Loss given default is measured using the aggregate loan loss provisions of the commercial banks.

financial margin, the probability of default, etc. The first column presents estimates based on the aggregate historical banking-sector data from the survey period, i.e., the second quarter of 2013. The second column presents survey-based financial fragility indicators, which are calculated using the estimation methodology outlined in the previous section. As discussed above, we have calibrated our model so that the survey-based exposure at default rate meets the actual historical nonperforming loan rate, which was 3.4 percent during the survey fieldwork period. The microdata-based loss given default rate (second column in table 1) is benchmarked against the aggregate historical loan loss provision rate (first column in table 1). The historical data come from the Bank of Estonia credit risk model.

The aggregate provision rates are higher than predicted by the microdata, and there are two possible reasons for that. First, provisions can also cover restructured loans, and second, the models used by commercial banks for provisioning may be more conservative than our definition of loss given default. In their internal credit risk models, banks usually use the estimate of the ready sale price of the real estate, which might be only 75 percent or 80 percent of the market value. Our definition of loss given default in the microdata is less conservative and is based on the full market value of the real estate. However, only a small number of loans have been written off in Estonia, even in the aftermath of the Great Recession, which suggests that banks have historically often been overprovisioning. These differences are evident in figure B.1 in appendix B, which illustrates the historical developments in the aggregate nonperforming loan rate, the loan loss provision rate, and the write-off rate in the household and corporate sectors.

According to the survey-based estimates presented in the second column of table 1, the share of households with a negative financial margin is 13.0 percent, which is similar in magnitude to the average figure for the euro area of 12.3 percent (Ampudia, van Vlokhoven, and Zochowski 2016).[10] This share of households with a negative financial margin corresponds to an average probability of default of 5.2 percent and exposure at default of 3.4 percent. Exposure at default is lower than the probability of default, which shows that households with a high probability of default have smaller debt stocks. This is in correspondence with the regularity that households with noncollateralized loans are usually financially more fragile. The value of loans exposed to default is EUR 165.9 million, which results in loss given default of EUR 20.6 million. The number of indebted households in the survey is 769, which corresponds to a share of indebted households of 30.5 percent.

[10] The percentage of households with a negative financial margin for the euro area is calculated using the Household Finance and Consumption Survey first-wave results and with basic consumption defined as 20 percent of the median income.

3.3 Households' Financial Fragility: Estimates Based on Survey and Register Data

This section compares the estimated financial fragility indicators that are based on the survey against the estimates based on data from administrative sources. The composite survey-based measures of household assets, liabilities, and income have been replaced one by one by the corresponding estimates from administrative sources to understand what effect their separate replacement has on financial fragility indicators. Table 2 presents the results. Column 2 in table 2 reports the results where the survey-based household income is replaced by administrative income data. The household incomes derived from administrative sources tend to be lower than those based on the survey, as the median yearly household disposable income is EUR 8,100 according to the survey and EUR 7,500 according to the administrative data (see appendix A). A possible reason for this difference is that the survey also covers nonreported income, at least in part. Alternatively, the difference may be caused by measurement error, i.e., income overreporting by survey participants. Lower estimated incomes for most of the percentiles from administrative sources raise the share of households with a negative financial margin and result in a higher exposure at default rate and larger losses for the banks.

The comparison of survey data with administrative records shows that low-income households overreport their incomes and high-income households underreport theirs in the Estonian Household Finance and Consumption Survey. The same pattern has also been detected in many earlier studies making such comparisons (e.g., Rodgers, Brown, and Duncan 1993; Bound et al. 1994; Pischke 1995). However, the turning point after which the survey-based estimates are below the register-based estimates for income occurs at a relatively high percentile of the income distribution in the Estonian Household Finance and Consumption Survey. Business income is underestimated in the register data, and the survey-based estimates show higher values for this income component. Since wages are the dominant form of income, however, they are the main source of the differences between the survey data and the administrative data. Notably, it is not the business income that generates a longer right-hand-side tail for the income distribution in the administrative

Table 2. Indicators of the Financial Fragility of Households and Potential Losses for Banks: Estimations Based on Survey and Administrative Sources

	Baseline from the Survey (1)	Replacing Income from Administrative Sources (2)	Replacing Debt from Administrative Sources (3)	Replacing Assets from Administrative Sources (4)	All Components from Administrative Sources (5)
Negative Financial Margin, %	13.0	17.0	10.5	13.0	15.6
Probability of Default, %	5.2	6.8	3.6	5.0	6.4
Exposure at Default, %	3.4	3.8	3.9	3.4	5.9
Mortgages, %	3.2	3.5	4.1	3.1	6.0
Noncollateralized Loans, %	8.8	12.0	1.5	9.2	5.0
Exposure at Default, Mln. EUR	165.9	186.1	220.6	164.3	336.2
Mortages, Mln. EUR	147.4	161.0	212.8	145.1	310.6
Noncollateralized Loans, Mln. EUR	18.5	25.1	7.8	19.3	25.6
Loss Given Default, %	0.4	0.8	0.5	1.5	1.1
Mortgages, %	0.0	0.3	0.4	1.1	0.7
Noncollateralized Loans, %	8.8	12.0	1.5	9.2	5.0
Loss Given Default, Mln. EUR	20.6	39.5	28.5	71.4	61.5
Mortgages, Mln. EUR	2.1	14.4	20.6	52.2	35.9
Noncollateralized Loans, Mln. EUR	18.5	25.1	7.8	19.3	25.6
No. of Obs.	769	769	944	769	944

Sources: Authors' calculations from Estonian Household Finance and Consumption Survey and administrative data.
Note: Indebted households are defined as households with collateralized debt and with consumer loans, not including leases, credit line overdrafts, and credit card debt.

data, but the wage income. Unlike other forms of income, state pensions are underestimated in the Estonian Household Finance and Consumption Survey. This finding is also in correspondence with the results of earlier studies, which generally indicate that income from state transfers tends to be underestimated in surveys (Bound, Brown, and Mathiowetz 2001).

The third column of table 2 reports the results when survey-based measures of debt servicing costs and the outstanding balance of loans are replaced by measures based on the data from administrative sources. The first outcome from this replacement is that the share of indebted households increases to 37.1 percent, indicating that the share of debt participation is underestimated in the survey. Further analysis by debt components shows that the participation in mortgages on the household main residence (HMR) is accurately predicted, while the participation in other real estate loans and consumer loans is underrepresented in the survey. The participation in these two debt types is substantially different in the survey data and the administrative data, as only 2.7 percent of households have other real estate loans according to the survey but 7.4 percent do according to the administrative data, while 13.2 percent of households have consumer loans according to the survey and 25.8 percent do according to the administrative data (see appendix A). The differences between the survey and administrative data for the median values of these debt types are not as substantial as those in the participation rates. One possible explanation for this tendency is respondent fatigue, as the data on household main residence mortgages are collected first, then the mortgages on other real estate, and then the consumer loans. An alternative possibility is response bias, as households may systematically underreport noncollateralized loans. There is also evidence from the United States that households tend to underreport the outstanding balance of consumer credit (Brown et al. 2011).

Replacing the debt items from the survey with the ones based on administrative sources adds households with relatively smaller loans and loan servicing costs to the group of indebted households, and the share of households with a negative financial margin decreases. However, as the total amount of loans covered increases, the exposure at default and loss given default after this replacement are higher than the results obtained with the survey-based measures.

Lastly, we replace the values of all real assets and liquid financial assets with the data from administrative sources. The probability of default declines slightly as a result, indicating that liquid assets are somewhat underreported in the survey. The estimated loss given default more than triples as a result of this replacement, which shows that the real estate values estimated from the data from administrative sources tend to be smaller than the survey-based values. Surprisingly, the mean and median values for the household main residence are higher in the survey data than in the administrative data, while the opposite holds for other real estate (see appendix A). However, as the participation rate is much higher in household main residences than in other real estate[11] and most of the outstanding mortgages are collateralized by the HMR, this component is behind the higher estimates of loss given default values based on the administrative data. The real estate prices from administrative sources are based on regional transaction prices in the survey period. They may overestimate or underestimate the value of the real estate in the region because of possible composition bias. Even so, the difference between the survey values and the administrative values is quite substantial, which indicates a possible overestimation of the real estate values based on the self-assessments by households in the survey. The tendency to overestimate the value of the real estate in surveys relative to administrative data (or average transaction prices) has also been found in some previous studies (see, e.g., Kiel and Zabel 1999 and Johansson and Klevmarken 2007). However, we are not aware of previous research indicating that the bias in the valuation of HMRs is positive in the surveys, while for other real estate it is negative, as we find on the basis of the Estonian Household Finance and Consumption Survey data set.

The fifth column of table 2 replaces all the components of the financial margin with measures derived using the data from administrative sources. This replacement results in a somewhat higher

[11]Like in debt components, there is also some evidence of survey fatigue if the survey answers for other real estate items are compared with administrative data. The questions about the HMR are asked first, followed by the questions on other real estate items. It is evident that survey respondents underreport owning other real estate. The participation rate for the HMR is 76 percent in the survey and 69 percent in the administrative data, while the participation in other real estate is 32 percent in the survey and 39 percent in the administrative data.

rate for the probability of default than the survey-based estimate, together with substantially higher losses for the banks, especially from mortgage loans. The values for loss given default estimated on the basis of data from administrative sources are closer to the aggregate historical loan loss provisions reported in the first column of table 1 than the survey-based baseline measures are. This indicates that the estimates of real estate values by the commercial banks are closer to the transaction prices from the register than to the self-estimation by households in the survey.

Our estimates show that the financial distress indicators based on survey data were lower than the measures based on administrative sources. This implies that if countries use different methods for collecting data, some using household interviews and others using registers, it can undermine the comparability of the results. As most similar studies use data from household interviews, it is likely that the level of household distress has in general been underestimated. The comparison of the survey and administrative data from Estonia indicates that households tend to overreport their incomes and the values of their HMRs and at the same time tend to underreport the outstanding balance of their loans. Although the finding for incomes is not entirely backed by previous studies making similar comparisons (e.g., Pischke 1995), earlier research has also shown that real assets, and real estate in particular, tend to be overreported and liabilities underreported in surveys (Johansson and Klevmarken 2007, Brown et al. 2011). This indicates that net wealth tends to be measured with an upward bias in the surveys. Therefore, risk assessments for the household sector that are based on survey data should in general underestimate the actual level of financial fragility relative to estimations that are based on administrative records.

4. Household Stress Tests: The Effect of Shocks on Financial Fragility

This section gives an overview of the stress tests based on the survey and administrative data, presenting the results of standardized individual shocks to three variables. The standardized shocks are defined as shocks of one, two, and three standard deviations in the base interest rate, unemployment, and real estate prices. The standard deviation is calculated on the basis of quarterly data covering the period 2004:Q1–2013:Q2,[12] which is the period from Estonia's accession to the EU until the time of the survey fieldwork. It is assumed that shocks occur instantaneously and that there is no feedback from the financial sector to the real economy.

4.1 The Interest Rate Shock

First, we assess the effect of base interest rate shocks on the financial fragility of households and on banking-sector losses from defaulting loans. The base interest rate shock is assumed to affect only mortgage loan payments with adjustable interest rates, while mortgage loan payments with fixed interest rates and noncollateralized loan payments are assumed to remain unaffected by this shock.[13] Mortgage loan payments with adjustable interest rates have two parts, the principal and the interest payments. The payments of the principal are unaffected by the shock, while the interest payments rise because of the higher base rate. It is also assumed that the base interest rate shock will affect the income earned from sight and savings accounts. However, the income from these sources is so small compared with other income sources that it has almost no effect on the financial margins of households.

As a result of increase in the base interest rate, the financial margin (FM_i) in equation (1) will decrease for an indebted household i

[12] The earlier years were excluded, as they were affected by structural changes caused by the transition from a planned economy to a market one. The market for mortgage loans was practically non-existent in the 1990s, and interest rates and the unemployment rate were substantially higher before EU accession than after it. See more discussion about housing and mortgage market developments in this period in Meriküll and Rõõm (2016).

[13] The Household Finance and Consumption Survey does not collect information on whether noncollateralized loans have flexible or adjustable interest rates, but it is known that most of these loans have fixed interest rates in Estonia. As much as 67 percent of the consumer loan stock issued by banks had fixed interest rates in the second quarter of 2013. These loans included leases for cars, which usually have adjustable interest rates and are not covered by the Household Finance and Consumption Survey (Bank of Estonia internal statistics from the Financial Stability Department). So the actual share of fixed interest rate loans among noncollateralized debt is even higher in the Household Finance and Consumption Survey data. Therefore we assume that interest rate shocks do not affect the servicing costs for noncollateralized loans, i.e., there is no pass-through of the base interest rate shock to noncollateralized loans.

because the increase in debt servicing costs (DP_i) is larger than the increase in disposable income (Y_i). A lower financial margin in turn leads to a higher probability of default depicted by equation (2) and consequently to a larger amount of loans being exposed at default (equation (3)) and higher loan losses given default (equation (4)).

The share of adjustable interest rate mortgages is 82 percent of the total mortgage stock in the Estonian Household Finance and Consumption Survey. This puts Estonia together with Luxembourg, Malta, the Netherlands, Portugal, Slovenia, and Spain in the group of countries with a relatively high share of adjustable-rate loans in the euro area (see Ampudia, van Vlokhoven, and Zochowski 2014 for estimates for the other euro-area countries). Consequently, the pass-through of this shock to the financial margin is relatively strong in Estonia.

The most common base interest rate in Estonia is the six-month EURIBOR. As much as 95 percent of all mortgage loans with adjustable interest rates are tied to this base rate in the Estonian Household Finance and Consumption Survey. The rest of the loans are tied to the EURIBOR rates with other durations or to the commercial banks' own base rates. As the other base rates also follow the dynamics of the six-month EURIBOR rate, all the adjustable interest rate mortgages are assumed to be affected by the shock to this variable.

The six-month EURIBOR was 0.318 percent at the time of the survey and its standard deviation for the post-EU-accession period was 1.413 percent. This means that shocks of one, two, and three standard deviations correspond to rises in the EURIBOR rate from 0.318 percent to 1.1731 percent, 3.144 percent, and 4.557 percent, respectively. Although a shock as large as three standard deviations should capture extreme developments, the highest shocked value of 4.557 percent is still 0.5 percentage point smaller than the highest value seen in the sample period, which was 5.176 percent in the second quarter of 2008. This indicates that the variation in the EURIBOR rate has been quite low.

The results of the shock to the EURIBOR rate are presented in table 3. For comparative purposes, we present the stress-testing results using the survey data and the register data. Shocks of one, two, and three standard deviations increase the share of households with a negative financial margin by about the same magnitude,

Table 3. The Effect of a Shock to the EURIBOR Interest Rate on the Financial Fragility of Households

	Survey Data				Register Data			
	Pre-stress, EURIBOR = 0.318%	**1 sd Shock, EURIBOR = 1.731%**	**2 sd Shock, EURIBOR = 3.144%**	**3 sd Shock, EURIBOR = 4.557%**	**Pre-stress, EURIBOR = 0.318%**	**1 sd Shock, EURIBOR = 1.731%**	**2 sd Shock, EURIBOR = 3.144%**	**3 sd Shock, EURIBOR = 4.557%**
Negative Financial Margin, %	13.0	13.8	14.6	15.3	15.6	17.0	17.7	18.8
Probability of Default, %	5.2	5.9	6.1	6.4	6.4	6.7	7.2	7.7
Exposure at Default, %	3.4	4.9	5.3	5.9	5.9	6.3	7.2	8.1
Mortgages, %	3.2	4.8	5.1	5.8	6.0	6.4	7.3	8.2
Noncollateralized Loans, %	8.8	8.8	8.8	8.8	5.0	5.1	5.3	5.8
Exposure at Default, Mln. EUR	165.9	240.6	258.3	289.7	336.2	355.2	407.1	456.4
Mortages, Mln. EUR	147.4	222.2	239.8	271.2	310.6	329.4	380.0	426.9
Noncollateralized Loans, Mln. EUR	18.5	18.5	18.5	18.5	25.6	25.7	27.1	29.5
Loss Given Default, %	0.4	0.4	0.4	0.4	1.1	1.1	1.2	1.4
Mortgages, %	0.0	0.0	0.0	0.1	0.7	0.7	0.8	1.0
Noncollateralized Loans, %	8.8	8.8	8.8	8.8	5.0	5.1	5.3	5.8
Loss Given Default, Mln. EUR	20.6	20.6	20.6	21.1	61.5	61.7	69.1	81.7
Mortgages, Mln. EUR	2.1	2.1	2.1	2.6	35.9	36.0	41.9	52.1
Noncollateralized Loans, Mln. EUR	18.5	18.5	18.5	18.5	25.6	25.7	27.1	29.5
No. of Obs.	769	769	769	769	944	944	944	944

Sources: Authors' calculations from Estonian Household Finance and Consumption Survey data.
Note: "sd" stands for standard deviation.

irrespective of whether these estimated effects are based on survey or register data. The EURIBOR shock of three standard deviations increases the share of households with a negative financial margin by 18 percent (from 13 percent to 15.3 percent) according to the survey data and by 21 percent (from 15.6 percent to 18.8 percent) according to the register data.

Exposure at default reacts to interest rate increases substantially more strongly according to the estimates based on the survey data than it does according to the register data. In response to the shock of three standard deviations, the EAD goes up to 5.9 percent from the pre-stress level of 3.4 percent, i.e., it rises by 74 percent. By contrast, the estimate of the EAD that is based on register data goes up from the pre-stress level of 5.9 percent to 8.1 percent in response to the shock of three standard deviations, which corresponds to an increase of 37 percent. This difference in the magnitudes of reaction is partially explainable by a lower base level of the EAD in the survey data than in the register data (3.4 percent versus 5.9 percent).

There are also differences between the survey and the register data in the reaction of the LGD to the interest rate shock. The potential losses from this shock are almost negligible according to the survey data but quite substantial according to the register-based estimates. The rise in the EURIBOR of three standard deviations increases the loss given default by only 2 percent according to the survey, while the LGD rate increases by 33 percent according to the register data. These divergent results are caused by biases in the survey estimates of liabilities and assets. The share of households with negative equity is much bigger according to the register data than according to the survey because households overestimate the value of real estate and underestimate the value of liabilities in the survey (see appendix A).

4.2 The Unemployment Shock

There are various ways to estimate the effect of an unemployment shock on the household financial margin. The simplest approaches assume equal unemployment risk across individuals (Johansson and Persson 2006, Herrala and Kaukko 2007), while more advanced approaches assume idiosyncratic shocks to the probability of unemployment, taking into account that individuals with different personal characteristics such as age, gender, and education have a different propensity for becoming unemployed (Albacete and Fessler 2010; Bilston, Johnson, and Read 2015; Galuščák, Hlaváč, and Jakubík 2016; Ampudia, van Vlokhoven, and Zochowski 2014; and Bańbuła et al. 2015). The last three of the papers cited take a step further and also model transitions from unemployment to employment on top of the probability of becoming unemployed.

Given our focus on the effects of adverse shocks, only the increase in the inflow from employment to unemployment is modeled in this paper. It is assumed that individuals who are unemployed at the time of the survey stay in unemployment after the shock. In addition, some individuals move from employment to unemployment, so that the increase in the unemployment rate matches the size of the shock. It is also assumed that the share of economically inactive people is unaffected by the shock. So our modeling of the unemployment shock assumes that the new and higher unemployment rate is caused by the change in one labor market flow, the flow from employment to unemployment, while other labor market flows remain unaltered.

These assumptions follow the logic of any labor market in recession where first hiring is cut and then separation increases because of adverse shocks (Davis and Haltiwanger 1999). The assumptions are also in line with the developments in the Estonian labor market during the Great Recession (see, e.g., Meriküll 2016). The unemployment rate mainly increased because of the high separation rate, while the hiring rate was very low throughout the crisis years. Despite the sluggish recovery of employment, job seekers did not switch from unemployment to inactivity, and the activity rate remained relatively stable over the boom, bust, and recovery.

The simulation of the unemployment shock is estimated using the approach taken by Albacete and Fessler (2010). Unlike in their analysis, unemployment is assessed at the individual level and not at the household level, and currently unemployed individuals are assumed to stay in unemployment in this paper. The shock is calculated in three steps. First, the predicted probability of each individual being unemployed is calculated using the logit model. Conventional regressors for the unemployment equation are used, such as gender, age, marriage, ethnicity, education, and region.[14]

Second, the constant term in the unemployment equation is manipulated to meet the new aggregate shock value of unemployment. Third, a random probability is drawn for each individual from a uniform distribution between zero and one. The model-based predicted probability of unemployment is compared with the random probability for each employed individual and if the predicted probability is larger than the random value, a switch from employment to unemployment is assigned for that person. Individuals who become unemployed are assigned new reduced gross incomes, which are equal to the previous gross wage income times the average replacement rate of 15 percent. The average replacement rate has been calculated using the crisis years of 2009 and 2010, which are taken as a good predictor of the replacement rate under a negative labor demand shock.[15]

The new household-level disposable income (Y_i) and financial margin (FM_i) are derived using equation (1), and the new values of the aggregate financial fragility indicators are calculated. This procedure is repeated 1,000 times using a Monte Carlo simulation,

[14] The marginal effects of the model are available from the authors upon request. Men, unmarried individuals, people of non-Estonian ethnicity, those with lower education, and those from Ida-Viru county have a higher probability of being unemployed.

[15] All workers who are involuntarily separated from work due to job destruction are subject to unemployment insurance in Estonia. The insurance benefit is 50 percent of the previous wage for the first three months and 40 percent for up to the next nine months dependent on the previous employment tenure. However, not all workers are eligible for the unemployment insurance. According to the Labour Force Survey, roughly 70 percent of workers who have moved from employment to unemployment within a year are registered with the unemployment insurance fund and, of these, only 50 percent receive unemployment insurance, while 25 percent receive unemployment benefit and 25 percent do not receive any transfers. These regularities held during the crisis years of 2009 and 2010, when most of the separations were due to job destruction. Given that we do not have enough detailed information on employment tenure and that not all the workers are eligible to receive unemployment insurance, we used as a replacement rate the average replacement rate of the crisis years from the Labour Force Survey. The income of an unemployed person is assumed to consist of an unemployment insurance payment, an unemployment benefit, and a training scholarship from the unemployment insurance fund. The severance payment is not covered by the Labour Force Survey. The size of the severance payment can be up to two months' salary.

and the effect of the unemployment shock is found as the average value of financial fragility indicators from these 1,000 replications. As a result of this shock, income will decrease for households where household members have switched from employment to unemployment, and in these households the disposable income and financial margin will decrease and lead to deterioration of the loans exposed and losses given default.

The unemployment rate is 10.9 percent in the Household Finance and Consumption Survey data, which is somewhat higher than the official estimates from the Labour Force Survey, which were 10.0 percent and 8.0 percent in the first and second quarters of 2013, respectively. The standard deviation of the official seasonally adjusted rate is 3.9 percent, which shows quite high variation for this variable. For example, the shock of three standard deviations would increase the sample unemployment rate to 22.8 percent, which is higher than the historical quarterly maximum since 2004 of 18.8 percent.

The effects of the unemployment shocks are presented in table 4. For comparative purposes, we first present the estimated effects that are based on survey data and then the results that are based on the register data. Shocks of one, two, and three standard deviations increase the share of households with a negative financial margin, as well as the probability of default and the exposure at default almost linearly. The share of households with a negative financial margin increases more strongly than it did in response to the interest rate shock. This shows that more households are affected by the unemployment shock. In response to the shock of three standard deviations, the share of households with a negative financial margin increases from the pre-stress level of 13 percent to 17.1 percent according to the survey data and from 15.6 percent to 21.7 percent according to the register data. The magnitude of this increase is 32 percent in the survey data and 39 percent in the register data.

The exposure at default is affected less strongly than in response to the interest rate shock, which indicates that the households affected by the unemployment shock usually have smaller loans than the households affected by the interest rate shock do. In response to the shock of three standard deviations, the EAD increases by 41 percent according to the survey data or by 29 percent according to the register data. Thus register-based estimates indicate smaller

Table 4. The Effect of the Unemployment Shock on the Financial Fragility of Households

	Survey Data				Register Data			
	Pre-stress, Unemployment Rate = 10.9%	**1 sd Shock, Unemployment Rate = 14.8%**	**2 sd Shock, Unemployment Rate = 18.8%**	**3 sd Shock, Unemployment Rate = 22.8%**	**Pre-stress, Unemployment Rate = 10.9%**	**1 sd Shock, Unemployment Rate = 14.8%**	**2 sd Shock, Unemployment Rate = 18.8%**	**3 sd Shock, Unemployment Rate = 22.8%**
Negative Financial Margin, %	13.0	14.3	15.7	17.1	15.6	17.3	19.5	21.4
Probability of Default, %	5.2	5.8	6.5	7.1	6.4	7.2	8.2	9.0
Exposure at Default, %	3.4	3.8	4.3	4.8	5.9	6.4	7.1	7.7
Mortgages, %	3.2	3.6	4.1	4.5	6.0	6.4	7.0	7.6
Noncollateralized Loans, %	8.8	9.4	9.9	10.5	5.0	6.2	7.3	8.3
Exposure at Default, Mln. EUR	165.9	186.0	210.3	232.9	336.2	365.2	401.0	434.8
Mortages, Mln. EUR	147.4	166.4	189.5	211.0	310.6	333.8	364.0	392.5
Noncollateralized Loans, Mln. EUR	18.5	19.6	20.8	21.9	25.6	31.4	37.0	42.3
Loss Given Default, %	0.4	0.5	0.5	0.6	1.1	1.3	1.5	1.7
Mortgages, %	0.0	0.1	0.1	0.1	0.7	0.8	0.9	1.0
Noncollateralized Loans, %	8.8	9.4	9.9	10.5	5.0	6.2	7.3	8.3
Loss Given Default, Mln. EUR	20.6	22.7	25.1	27.2	61.5	73.0	85.1	95.9
Mortgages, Mln. EUR	2.1	3.1	4.3	5.4	35.9	41.5	48.1	53.5
Noncollateralized Loans, Mln. EUR	18.5	19.6	20.8	21.9	25.6	31.4	37.0	42.3
No. of Obs.	769	769	769	769	944	944	944	944

Sources: Authors' calculations from Estonian Household Finance and Consumption Survey data.
Note: "sd" stands for standard deviation.

sensitivity to unemployment shocks for this variable than survey-based estimates do.

At the same time, the potential losses for the banks from the unemployment shock are larger than those from the interest rate shock. The LGD rate increases by 50 percent in response to the shock of three standard deviations according to the survey data and by 55 percent according to the register data. Evidently, households that are more strongly affected by the unemployment shock have higher loan-to-value ratios than households that are affected more by the interest rate shock, which results in a larger exposure of banks to the unemployment shocks.

4.3 The Real Estate Price Shock

The real estate price shock does not affect the financial margin of households, since the financial margin is calculated using flow variables related to income and expenditures and does not depend on the value of assets. This means that the probability of default and exposure at default do not depend on the real estate price shock either and this shock affects only loss given default, i.e., equations (1)–(3) will not be affected. A fall in real estate prices increases loan-to-value ratios and the number of households with negative equity. The losses given default that are calculated using equation (4) will increase because the value of assets that the bank can liquidate in the event of a default (W_i) will decrease and the number of households "underwater" (c_i) will increase.

Real estate prices have historically been very volatile in Estonia, which experienced a very strong boom-and-bust cycle in house prices, culminating before the Great Recession. House prices dropped by 50 percent between 2007 and 2009 in Estonia, and while this was followed by a recovery in the real estate market, house prices were still only at 70 percent of their highest historical value at the time of the survey. The volatile development of real estate prices is reflected in the high standard deviation in this variable. One standard deviation in the real estate price index corresponds to a fall of 24.4 percent in prices. This is much higher than in other euro-area countries and is even substantially higher than in Spain, the country with the highest standard deviation in the study by Ampudia, van Vlokhoven, and Zochowski (2016), where it was 14.3 percent. Given the sizable standard deviation in Estonia, a shock of three standard deviations cannot be considered realistic. The two-standard-deviation shock of 48.8 percent corresponds to the decline in real estate prices during the Great Recession.

The results of the real estate price shock are presented in table 5. For comparative purposes, we first present the estimated effects that are based on survey data and then the results that are based on the register data. Although it is only the losses from mortgage loans that are affected by these shocks, they have a strong adverse effect on loan losses. Losses given default increase by almost five times (from 0.4 percent to 1.9 percent) according to the survey data and by about 220 percent (from 1.1 percent to 3.5 percent) according to the register data. There is also a strong nonlinearity in this reaction, as the reaction in losses is much stronger to larger shocks. The takeaway from this nonlinearity is that the bulk of households with a negative financial margin have high loan-to-value ratios and the deterioration in real estate prices drives them quickly into negative equity.

As a result of the shock of three standard deviations, the expected losses for banks are more than four times larger than the pre-stress level and reach EUR 92 million according to the survey data and EUR 197 million according to the register data. Although this is a strong increase from the baseline scenario, the estimated level of losses according to the survey is similar in magnitude to the aggregate quarterly profits of the banking sector, which were about EUR 90 million per quarter in 2011–13. The estimated level of losses is about twice this value in the worst-case scenario according to the register data. Thus, even in the case of an extreme decline of 73 percent in real estate prices, the estimated losses from loan defaults by the household sector are relatively easy for the commercial banks to absorb.

The comparison of the effects of the three shocks implies that interest rate shocks have the mildest effect, followed by unemployment shocks. Shocks to real estate prices do not affect the ability of households to service their loans but have the strongest effect on the estimated loan losses of the banks.

The finding that the financial fragility of households is more sensitive to unemployment shocks than to interest rate shocks is

Table 5. The Effect of the Shock to Real Estate Prices on the Financial Fragility of Households

	Survey Data				Register Data			
	Pre-stress	**1 sd Shock, Real Estate Price Decrease = 24.4%**	**2 sd Shock, Real Estate Price Decrease = 48.8%**	**3 sd Shock, Real Estate Price Decrease = 73.2%**	**Pre-stress**	**1 sd Shock, Real Estate Price Decrease = 24.4%**	**2 sd Shock, Real Estate Price Decrease = 48.8%**	**3 sd Shock, Real Estate Price Decrease = 73.2%**
Negative Financial Margin, %	13.0	13.0	13.0	13.0	15.6	15.6	15.6	15.6
Probability of Default, %	5.2	5.2	5.2	5.2	6.4	6.4	6.4	6.4
Exposure at Default, %	3.4	3.4	3.4	3.4	5.9	5.9	5.9	5.9
Mortgages, %	3.2	3.2	3.2	3.2	6.0	6.0	6.0	6.0
Noncollateralized Loans, %	8.8	8.8	8.8	8.8	5.0	5.0	5.0	5.0
Exposure at Default, Mln. EUR	165.9	165.9	165.9	165.9	336.2	336.2	336.2	336.2
Mortages, Mln. EUR	147.4	147.4	147.4	147.4	310.6	310.6	310.6	310.6
Noncollateralized Loans, Mln. EUR	18.5	18.5	18.5	18.5	25.6	25.6	25.6	25.6
Loss Given Default, %	0.4	0.6	1.1	1.9	1.1	1.4	2.1	3.5
Mortgages, %	0.0	0.3	0.8	1.6	0.7	1.1	1.8	3.3
Noncollateralized Loans, %	8.8	8.8	8.8	8.8	5.0	5.0	5.0	5.0
Loss Given Default, Mln. EUR	20.6	31.0	55.3	91.6	61.5	80.7	120.8	197.1
Mortgages, Mln. EUR	2.1	12.6	36.8	73.1	35.9	55.0	95.1	171.5
Noncollateralized Loans, Mln. EUR	18.5	18.5	18.5	18.5	25.6	25.6	25.6	25.6
No. of Obs.	769	769	769	769	944	944	944	944

Sources: Authors' calculations from Estonian Household Finance and Consumption Survey data.
Note: "sd" stands for standard deviation.

similar to the earlier results for other Central and Eastern European countries, which also mostly indicate that unemployment rate shocks have stronger adverse effects than interest rate shocks do (see Galuščák, Hlaváč, and Jakubík 2016 and Hóllo and Papp 2007). It is in contrast with the findings for some Western European and Nordic countries, which typically show the opposite pattern (e.g., Johanssson and Persson 2006, Herrala and Kauko 2007, and Albacete and Fessler 2010). Unemployment shocks have a stronger effect on households' loan servicing ability in Central and Eastern European countries, including Estonia, because the income replacement rate for those who become unemployed is lower in this region.

However, the most harmful shock for financial stability in Estonia is the decline in real estate prices, which leads to the largest losses for the banks. This is related to the fast and substantial debt accumulation of Estonian households and the historically volatile real estate prices, so that the loan-to-value ratios of mortgage loans are high and the simulated shocks of one, two, and three standard deviations are very sizable and have a strong effect on the value of real estate assets.

4.4 The Effect of Standardized Shocks across Households with Different Characteristics

The previous subsections describing the effect of various shocks on the financial fragility of households showed the aggregate reaction to the deterioration in each shocked variable, but the discussion of aggregate reaction did not say much about the heterogeneous reaction of households. This subsection will review which households are more vulnerable to shocks and which households are responsible for most of the loan losses that occur because of the shocks. The households are grouped by four characteristics: net income, net wealth, age of the household reference person,[16] and household size.

[16]The household reference person is defined following the Canberra definition (United Nations Economic Commission for Europe 2011). See Meriküll and Rõõm (2016) for a description of its derivation.

Figure 1. Variation in the Probability of Default (%) across Households with Different Characteristics and in Response to Different Shocks

Source: Authors' calculations from the Estonian Household Finance and Consumption Survey data.
Notes: "sd" stands for standard deviation. The value for households in the 75+ age group is not reported, as there were fewer than 20 such indebted households in the sample.

The financial vulnerability of households is best captured by the probability of default. Figure 1 shows the variation in the average value of the probability of default across the household characteristics listed above and for four scenarios: pre-stress, interest rate shock of one standard deviation, unemployment shock of one standard deviation, and real estate price shock of one standard deviation. Financial vulnerability is highest for low-income households but is also above average for small households, low-net-wealth households, and households with a middle-aged reference person. In general, the survey and register data point to the same groups of vulnerable households. The variation in the probability of default is strongest across income groups, ranging from 40 percent for the lowest quintile to near-zero values for the upper three quintiles.

There are also some differences between the survey and register data in the sensitivity of the probability of default to various shocks.

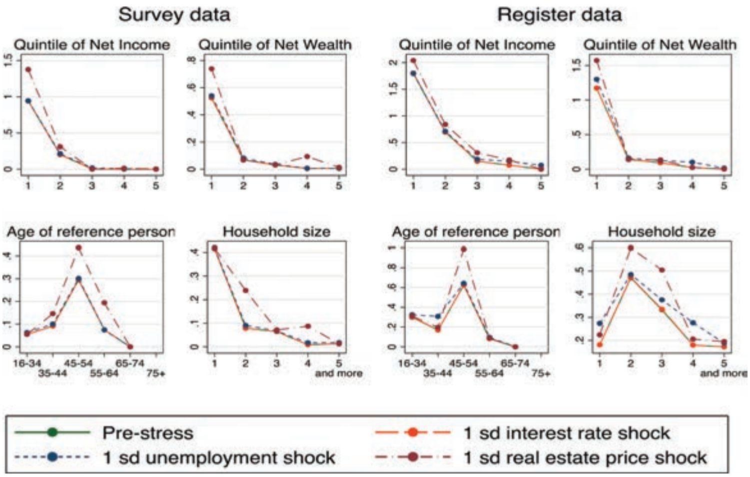
First, the household sector is relatively more vulnerable to the interest rate shock according to the survey and relatively more vulnerable to the unemployment rate shock according to the register. The possible reason for this is that noncollateralized loans are more important in the register (see appendix A). Since these loans are unaffected by the interest rate shock, they make the register-data-based estimates less responsive to this shock. The strong responsiveness to the unemployment shock in the register data can be explained by the lower estimated income from the register (see appendix A), which makes the financial margin more sensitive to income shocks from household members becoming unemployed. The real estate price shock has no effect on the probability of default, since it does not affect the flow variables (income and expenditures) that drive the probability.

The second difference between the survey and register data is the more dispersed distribution of the probability of default across household groups in the survey than in the register. According to the survey data, the probability of default is high for the first income quintile but close to zero for the upper three quintiles. According to the register data, the default risk is also quite high in the second and third income quintiles. The probability of default is also more dispersed across the wealth quintiles according to the register data than according to the survey data.

Third, the reaction to shocks is also more concentrated according to the survey than according to the register data. The effect of the interest rate shock is clearly concentrated to households in the 45 to 54 age group in the survey, while the reaction is more equally distributed across age groups in the register. It is hard to tell why the distributions of probability of default differ like this in the survey and in the register. It may be that the survey data capture extreme values or tails of the distribution better than the register or it could be that some population groups systematically misreport in the survey.

The differences in the effect of shocks across household characteristics are more pronounced in the monetary value of loss given default than in the probability of default. Loss given default is a better characterization of the risks for lenders, as a high probability of default does not necessarily imply high risks for the financial sector if the amounts of debt involved are small. Figure 2 presents loss given default in thousands of euros per household. Loss given

Figure 2. Variation in Loss Given Default across Households with Different Characteristics and in Response to Different Shocks, in Thousands of EUR per Household



Source: Authors' calculations from the Estonian Household Finance and Consumption Survey data.
Notes: "sd" stands for standard deviation. The value for households in the 75+ age group is not reported, as there were fewer than 20 such indebted households in the sample.

default is more concentrated in specific household groups than the probability of default is. Households from the lowest income quintile, from the lowest net wealth quintile, from the 45–54 age group, and single-person households are responsible for the majority of losses. The real estate price shock leads to larger losses for the banking sector than the other standardized shocks do, and the effect of this shock is the strongest for households in the lowest income and net wealth quintiles.

Losses for the banks are again distributed with greater dispersion in the register than in the survey, while both of the data sources point to the same vulnerable segments of households. The results are slightly different in the survey and register across household size, as single-member households bear the largest risks for banks according to the survey and two-member households do according to the register. The two-member households apparently have negative equity much more frequently in the register than in the survey.

However, the main difference between the survey and register data in the loss given default is the size of the losses—the register data predict substantially larger losses than the survey data do. There are two main reasons for this. First, the participation in debt and the share of households with negative equity are higher according to the register data than according to the survey (see appendix A). Second, the share of households with negative net value of housing equity is larger in the register data than in the survey since households tend to overvalue their real estate holdings in the survey. The second reason has the dominant role in driving the result that losses from default are higher according to the register data. The differences in debt participation rates account for only one-fifth of the differences in losses given default, while the majority of differences are caused by the share of households with negative equity being larger in the register data.

5. Conclusion

This paper assesses the financial fragility of the household sector using data for the same individuals from household interviews and administrative registers. We employ a set of stress-testing exercises in which the probability of default on loans is evaluated on the basis of households' financial margins and net liquid financial assets. The analysis employs microlevel data from the Estonian Household Finance and Consumption Survey.

The comparison of the survey and administrative data from Estonia indicates that households tend to overreport their incomes and the values of their real estate and at the same time tend to underreport their outstanding balance of loans in surveys. Although the finding for incomes is mostly not backed by previous studies making similar comparisons, earlier research has also shown that real assets tend to be overreported and liabilities underreported in surveys. This implies that net wealth tends to be measured with an upward bias in surveys. Therefore, risk assessments for the household sector that are based on the survey data are likely to underestimate the level of financial fragility, relative to estimations that are based on administrative records.

We assessed the financial risks of the household sector using three indicators: households' probability of default, exposure of loans to default, and banking-sector loan losses from defaulting household loans. Because of the biases described in the previous paragraph, the estimated financial fragility indicators tended to be lower when the estimates were based on the survey data, relative to the estimates based on the administrative data. These differences were more substantial for banks' losses given default than for households' probability of default. The probable reason for this is that the income estimates from the survey and register data are more similar than the estimates of assets and liabilities are.

After the assessment of the financial fragility indicators, we performed stress tests to evaluate how sensitive the financial vulnerability of the household sector is to adverse shocks. The stress-test elasticities of household default rates and banking-sector loan losses were assessed separately for three standardized negative macroeconomic shocks: a rise in interest rates, an increase in the unemployment rate, and a fall in real estate prices. The comparison of the effects of these three shocks implies that interest rate shocks have the mildest effect, followed by the unemployment shocks. Shocks to real estate prices do not affect the ability of households to service their loans, but they have the strongest effects on the estimated loan losses of the banks.

The finding that the financial fragility of households is more sensitive to unemployment shocks than to interest rate shocks in Estonia is similar to the earlier results for other Central and Eastern European countries, which also mostly indicate that unemployment rate shocks have stronger adverse effects than interest rate shocks do (e.g., Galuščák, Hlaváč, and Jakubík 2016, Hóllo and Papp 2007). It is in contrast with the findings for some Western European and Nordic countries, which typically show the opposite pattern (e.g., Johanssson and Persson 2006, Herrala and Kauko 2007, and Albacete and Fessler 2010). Unemployment shocks have a stronger effect on the ability of households to service their loans in Central and Eastern European countries, including Estonia, because the income replacement rate for people who become unemployed is lower in this region.

The most harmful shock for financial stability in Estonia is the decline in real estate prices, which leads to the largest losses for the banks. This is related to the fast and substantial debt accumulation of Estonian households during the boom years preceding the Great Recession and the historically volatile real estate prices, so that the loan-to-value ratios of mortgage loans were high in the survey year of 2013 and the simulated shocks of one, two, and three standard deviations had a strong negative effect on the value of real estate assets.

The stress-testing exercises were performed separately on the survey data and on the register data. The effect of the adverse shocks using the survey data was similar in percentage terms to that based on the register data. However, since the pre-stress level of households' financial fragility was larger according to the register data, the resulting effects on the levels of financial indicators from the shocks were substantially larger in the register data. This applies especially to losses given default, which were up to four times larger according to the register data than according to the survey data. This result questions the comparability of stress-testing results from survey and administrative data and implies that the risks to the banking sector may be underestimated if the assessments are based on household surveys.

Next, we studied the heterogeneity in households' financial fragility. This was done by grouping households on the basis of different characteristics and evaluating the financial fragility indicators and reaction to shocks across these household groups. The households were grouped by four characteristics: net income, net wealth, age of the household reference person, and household size. Financial vulnerability varied the most across income quintiles, and was the highest for low-income households. The probability of default ranged from 40 percent for the lowest quintile to near-zero values for the upper quintiles. A similar pattern could be observed for loss given default, which was substantial for the lowest income quintile and fell to almost zero level for the upper three quintiles. The sensitivity to shocks in the case of loss given default was also the highest for low-income and low-net-wealth households. This implies that most of the financial risks from household loans are borne by low-income households in Estonia.

The heterogeneity in the financial vulnerability indicators of households was studied separately using the survey data and register data. In general, the survey and the register data pointed to the same groups of vulnerable households, as both data sources indicated that low-income households were the most vulnerable. The distribution of risks tended to be more concentrated to low-income and low-net-wealth households according to the survey data than according to the register data, but these differences in concentration were relatively small.

The stress-testing framework used in this paper can be extended to assess the effect of changes in other variables, such as taxes or consumer prices. In the future, similar stress-testing exercises can be repeated using updated versions of the Household Finance and Consumption Survey data. It will be possible to employ either the future waves of the survey data, which will be collected at three-year intervals, or simulated data, which can be compiled with a higher frequency. Another avenue for further research is to perform stress tests on survey and administrative data for other countries too. It is important to understand differences in these data sources, as there is an increasing tendency to use administrative data for assessing financial risks.

Appendix A. Descriptive Statistics: Survey Data and Register Data

Table A.1. Descriptive Statistics: Survey Data and Register Data

	Obs.	Share in Population	Min.	P5	P25	P50	P75	P95	Max.	Mean	Total (Mln. EUR)
Assets											
Value of Household's Main Residence (HMR)											
Survey	1,778	0.765	200	4,908	21,148	44,850	75,386	200,000	1,405,385	69,152	30,240
Register	1,636	0.693	1,946	7,223	18,183	39,976	62,631	144,139	546,449	50,614	20,049
Value of Other Real Estate Property											
Survey	834	0.320	1	1,529	9,913	27,197	60,000	200,000	1,105,264	57,434	10,514
Register	1,011	0.394	13	3,054	16,133	36,293	77,183	222,743	1,932,445	69,217	15,589
Total Value of Real Estate Property											
Survey	1,870	0.805	200	5,000	25,000	50,048	100,000	289,343	1,586,472	88,480	40,754
Register	1,803	0.767	371	8,014	23,912	49,462	96,380	242,998	2,364,311	81,206	35,638
Net Liquid Assets											
Survey	2,220	1.000	0	0	60	1,162	6,152	35,420	1,240,302	8,826	5,047
Register	2,220	1.000	0	0	20	1,250	7,191	38,315	3,956,984	10,535	6,024
Liabilities											
Outstanding Balance of HMR Mortgages											
Survey	509	0.187	260	2,092	10,000	27,623	53,500	111,800	300,000	39,002	4,164
Register	476	0.177	276	2,220	12,059	28,150	55,248	116,272	295,999	40,042	4,060
Outstanding Balance of Other Mortgages											
Survey	79	0.027	100	809	8,080	21,758	52,000	92,000	170,000	32,140	501
Register	202	0.074	116	2,102	7,446	16,680	33,319	82,537	166,887	25,509	1,075
Outstanding Balance of Noncollateralized Loans											
Survey	299	0.132	9	122	600	1,396	3,317	8,600	56,000	2,768	209
Register	634	0.258	0	135	839	1,786	4,055	10,469	58,987	3,444	508

(*continued*)

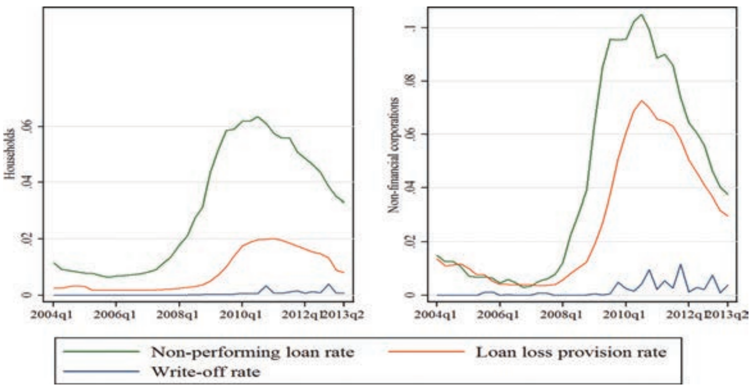
Table A.1. (Continued)

	Obs.	Share in Population	Min.	P5	P25	P50	P75	P95	Max.	Mean	Total (Mln. EUR)
Total Outstanding Balance of Mortgages and Noncollateralized Loans											
Survey	769	0.305	10	400	2,449	12,000	40,500	100,000	300,000	27,960	4,875
Register	944	0.371	65	465	2,525	10,469	36,543	102,600	321,263	26,671	5,665
Negative Equity ($(D_i^M - W_i^M) c_i^M i$ in Equation (1))											
Survey	29	0.010	612	923	4,883	8,123	23,209	76,106	104,032	19,004	111
Register	132	0.047	13	714	6,543	13,757	26,795	90,437	200,210	24,622	662
Total Household Disposable Income											
Survey	2,196	0.982	145	2,336	4,045	8,073	14,901	30,873	363,110	11,690	6,567
Register	2,131	0.941	–19,163	1,828	3,742	7,465	13,857	30,147	9,262,065	12,721	6,844
Debt Servicing Costs											
Payments for HMR Mortgages											
Survey	509	0.187	0	72	135	210	320	650	2,900	273	29
Register	476	0.177	26	75	128	209	354	708	2,823	278	28
Payments for Other Property Mortgages											
Survey	79	0.027	7	50	111	198	292	555	958	223	3
Register	202	0.074	19	55	104	153	266	519	19,243	350	15
Payments for Noncollateralized Debt											
Survey	299	0.132	0	8	38	68	123	329	1,403	107	8
Register	634	0.258	3	17	43	75	137	276	468	102	14
Payments for Household's Total Debt											
Survey	769	0.305	0	22	95	180	300	604	2,900	233	41
Register	944	0.371	7	31	93	174	319	643	19,243	266	56
Financial Margin (See Equation (1)), Monthly											
Survey	769	0.305	–1,984	–167	227	652	1,247	2,663	11,672	902	157
Register	944	0.371	–17,798	–319	165	580	1,108	2,464	14,232	779	166

Sources: Authors' calculations from Estonian Household Finance and Consumption Survey and administrative data.

Appendix B. Trends in Loan Quality and Banking-Sector Losses in Estonia

Figure B.1. Nonperforming Loan Rates, Loan Loss Provision Rates, and Write-off Rates, 2004:Q1–2013:Q2



Source: The Bank of Estonia statistics table 3.3.11 for the nonperforming loans, and the Bank of Estonia credit risk model for loan loss provisions and write-offs.
Note: The nonperforming loan rate is based on loans past due by more than 30 days.

References

Albacete, N., and P. Fessler. 2010. “Stress Testing Austrian Households.” In *Financial Stability Report 19*, June 2010, 72–91. Oesterreichische Nationalbank.

Ampudia, M., H. van Vlokhoven, and D. Zochowski. 2014. “Financial Fragility of Euro Area Households.” ECB Working Paper No. 1737.

———. 2016. “Financial Fragility of Euro Area Households.” *Journal of Financial Stability* 27 (December): 250–62.

Andersen, A., L., A. M. Christensen, C. Duus, and R. Kaarup. 2013. “Danish Families’ Robustness, Variable Rates and Deferred Amortisation.” *Monetary Review* (Danmarks Nationalbank) 4th Quarter (Part 2): 1–52.

Bańbuła, P., A. Kotuła, J. Przeworska, and P. Strzelecki. 2015. "Which Households Are Really Financially Distressed: How MICRO-data Could Inform the MACRO-prudential Policy." Paper presented at the BIS Irving Fisher Committee on Central Bank Statistics Workshop, Warsaw, December 14–15.

Bilston, T., R. Johnson, and M. Read. 2015. "Stress Testing the Australian Household Sector Using the HILDA Survey." Research Discussion Paper No 2015-01, Reserve Bank of Australia.

Bound, J., C. Brown, G. J. Duncan, and W. L. Rodgers. 1994. "Evidence on the Validity of Cross-sectional and Longitudinal Labor Market Data." *Journal of Labor Economics* 12 (3, July): 345–68.

Bound, J., C. Brown, and N. Mathiowetz. 2001. "Measurement Error in Survey Data." In *Handbook of Econometrics*, Vol. 5, ed. J. J. Heckman and E. Leamer, 3705–3843 (chapter 59). Amsterdam: Elsevier.

Bricker, J., A. Henriques, J. Krimmel, and J. Sabelhaus. 2016. "Measuring Income and Wealth at the Top Using Administrative and Survey Data." *Brookings Papers on Economic Activity* (Spring): 261–312.

Brown, M., A. Haughwout, D. Lee, and W. van der Klaauw. 2011. "Do We Know What We Owe? A Comparison of Borrower- and Lender-Reported Consumer Debt." Staff Report No. 523, Federal Reserve Bank of New York.

Chen, X., H. Hong, and D. Nekipelov. 2011. "Nonlinear Models of Measurement Errors." *Journal of Economic Literature* 49 (4): 901–37.

Davis, S., J., and J. Haltiwanger. 1999. "Gross Job Flows." In *Handbook of Labor Economics*. ed. O. C. Ashenfelter and D. Card, 2711–2805 (chapter 41). Amsterdam: North-Holland.

Duncan, G. J., and D. H. Hill. 1985. "An Investigation of the Extent and Consequences of Measurement Error in Labor-Economic Survey Data." *Journal of Labor Economics* 3 (4): 508–32.

Eurosystem Household Finance and Consumption Network. 2013a. "The Eurosystem Household Finance and Consumption Survey: Methodological Report for the First Wave." ECB Statistics Paper No. 1 (April).

———. 2013b. "The Eurosystem Household Finance and Consumption Survey: Results from the First Wave." ECB Statistics Paper No. 2 (April).

———. 2016. "The Household Finance and Consumption Survey: Methodological Report for the Second Wave." ECB Statistics Paper No. 17 (December).

Faruqui, U., X. Liu, and T. Roberts. 2012. "An Improved Framework for Assessing the Risks Arising from Elevated Household Debt." *Financial System Review* (Bank of Canada) (June): 51–57.

Galuščák, K., P. Hlaváč, and P. Jakubík. 2016. "Household Resilience to Adverse Macroeconomic Shocks: Evidence from Czech Microdata." *International Review of Applied Economics* 30 (3): 377–402.

Herrala, R., and K. Kauko. 2007. "Household Loan Loss Risk in Finland: Estimations and Simulations with Micro Data." Research Discussion Paper No. 5/2007, Bank of Finland.

Holló, D., and M. Papp. 2007. "Assessing Household Credit Risk: Evidence from a Household Survey." Occasional Paper No. 70/2007, Magyar Nemzeti Bank.

Israel, J.-M., V. Damia, R. Bonci, and G. Watfe. 2017. "The Analytical Credit Dataset: A Magnifying Glass for Analysing Credit in the Euro Area." Occasional Paper No. 187 (April), European Central Bank.

Johansson, F., and A. Klevmarken. 2007. "Comparing Register and Survey Wealth Data." Department of Economics, Uppsala University.

Johansson, M. W., and M. Persson. 2006. "Swedish Households' Indebtedness and Ability to Pay — A Household Level Study." *Economic Review* (Sveriges Riksbank) (2006:3): 24–41.

Kapteyn, A., and J. Y. Ypma. 2007. "Measurement Error and Misclassification: A Comparison of Survey and Administrative Data." *Journal of Labor Economics* 25 (3): 513–51.

Kiel, K., A., and J. E. Zabel. 1999. "The Accuracy of Owner-provided House Values: The 1978-1991 American Housing Survey." *Real Estate Economics* 27 (2): 263–98.

Kukk, M. 2019. "Debt Repayment Problems: Short-term and Long-term Implications for Spending." *Review of Economics of the Household* 17 (2): 715–40.

Kukk, M., and K. Staehr. 2014. "Income Underreporting by Households with Business Income: Evidence from Estonia." *Post-Communist Economies* 26 (2): 257–76.

Martinez, F., R. Cifuentes, C. Madeira, and R. Poblete-Cazenave. 2013. "Measurement of Household Financial Risk with the Survey of Household Finances." Documento de Trabajo No. 682, Banco Central de Chile.

Meriküll, J. 2016. "Labor Market Transitions during the Great Recession in Estonia." In *Labor Migration, EU Enlargement, and the Great Recession,* ed. M. Kahanec and K. F. Zimmermann, 347–65. Berlin/Heidelberg: Springer Verlag.

Meriküll, J., and T. Rõõm. 2016. "The Assets, Liabilities and Wealth of Estonian Households: Results of the Household Finance and Consumption Survey." Occasional Paper No. 1/2016, Eesti Pank.

Mian, A., and A. Sufi. 2010. "The Great Recession: Lessons from Microeconomic Data." *American Economic Review* 100 (2): 51–56. Papers and Proceedings of the One Hundred Twenty Second Annual Meeting of the American Economic Association.

Michelangeli, V., and M. Pietrunti. 2014. "A Microsimulation Model to Evaluate Italian Households' Financial Vulnerability." Occasional Paper No. 225, Banca d'Italia.

Pischke, J.-S. 1995. "Measurement Error and Earnings Dynamics: Some Estimates from the PSID Validation Study." *Journal of Business and Economic Statistics* 13 (3): 305–14.

Putnins, T., and A. Sauka. 2015. "Measuring the Shadow Economy Using Company Managers." *Journal of Comparative Economics* 43 (2): 471–90.

Rodgers, W. L., C. Brown, and G. J. Duncan. 1993. "Errors in Survey Reports of Earnings, Hours Worked, and Hourly Wages." *Journal of the American Statistical Association* 88 (424, December): 1208–18.

United Nations Economic Commission for Europe. 2011. *Canberra Group Handbook on Household Income Statistics.* 2nd edition. Geneva: United Nations Economic Commission for Europe.

Bank Capital: A Seawall Approach*

Jihad Dagher,[a] Giovanni Dell'Ariccia,[a,c] Luc Laeven,[b,c]
Lev Ratnovski,[a,b] and Hui Tong[a]
[a]International Monetary Fund
[b]European Central Bank
[c]CEPR

We find that bank capital in the range of 15–23 percent of risk-weighted assets would have been sufficient to absorb losses in the vast majority of historic banking crises in advanced economies. Further capital increases would have had only marginal effects on preventing additional crises. Appropriate capital requirements may be below this range, as banks tend to hold capital in excess of regulatory minimums, and other bail-in-able instruments can contribute to banks' loss-absorption capacity. While the long-term social costs associated with this level of capital appear acceptable, the short-term costs of transitioning to higher bank capital may be substantial, which calls for a careful timing of such transition.

JEL Code: G20.

1. Introduction

A large part of the post-crisis policy debate has focused on the appropriate levels of bank capital. Proponents of stricter regulation emphasize the risks and inefficiencies associated with high leverage and point to the exorbitant costs of the crisis (Admati and Hellwig 2014). Opponents of higher capital requirements believe that

*This paper is largely based on an IMF Staff Discussion Note, "Benefits and Costs of Bank Capital." We thank Harrison Hong (the editor), an anonymous referee, Olivier Blanchard, Charles Calomiris, Maurice Obstfeld, and Thomas Philippon for useful comments, and Nicola Babarcich, Francesca Caucci, Jae Chung, Paola Ganum, Huy Nguyen, and Yangfan Sun for excellent research assistance. The paper represents the views of the authors, and not those of the IMF, ECB, or their respective management or executive boards. Authors' e-mails: jdagher@imf.org; gdellariccia@imf.org; luc.laeven@ecb.int; lev.ratnovski@ecb.int; htong@imf.org.

Table 1. Basel I, Basel II, and Basel III Capital Requirements (percent of risk-weighted assets)

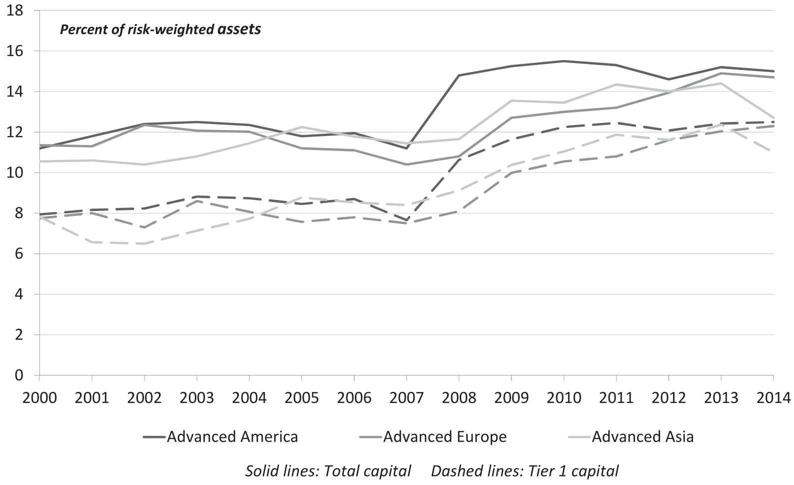
	Basel I	Basel II	Basel III[a]
Quantity of Capital			
Minimum Total Capital	8.0	8.0	8.0
Capital Conservation Buffer[b]	NA	NA	2.5
Minimum Total Capital Plus Conservation Buffer	NA	NA	10.5
Countercyclical Buffer[b]	NA	NA	0–2.5
Global Systemically Important Banks (G-SIB) Surcharge[b]	NA	NA	1–2.5
Minimum Total Capital Plus Conservation Buffer, Countercyclical Buffer, and G-SIB Charge	8.0	8.0	11.5–15.5
Leverage Ratio[c]	NA	NA	3.0
Quality of Capital			
Minimum Common Equity Capital[d]	NA	NA	4.5
Minimum Tier 1 Capital	4.0	4.0	6.0
Hybrid Capital Instruments with Incentive to Redeem[e]	Eligible	Eligible	Not Eligible

Source: Bank for International Settlements (2011, 2013).
[a]Effective as of 2019. [b]Consisting of tangible common equity. [c]Ratio of tier 1 capital to total assets. [d]Goodwill and deferred tax assets are to be deducted in the calculation of common equity tier 1 capital. [e]Hybrid capital instruments with an incentive to redeem through features such as step-up clauses, which under Basel II counted toward tier 2 capital and up to 15 percent of the tier 1 capital base, are no longer eligible as capital. Under Basel III only dated subordinated debt is deemed tier 2 capital.

these would increase the cost of financial intermediation and hinder economic activity (Institute of International Finance 2010). Some also caution that tighter regulation might push intermediation out of the banking system and into unregulated entities, possibly increasing systemic risk.

According to the prevalent view, pre-crisis bank capitalization proved insufficient to absorb crisis-related bank losses. In response, Basel III raised minimum bank capital requirements from 8 percent to up to 15.5 percent of risk-weighted assets, when all surcharges are activated (table 1). It also introduced a leverage ratio requirement

Figure 1. Tier 1 and Total Capital Ratios for Large Global Banks since 2000



and raised the quality of capital by tightening eligibility requirements for instruments included in the numerator of regulatory ratios (including requiring a larger fraction of regulatory capital to consist of tangible common equity). Some jurisdictions opted for even higher standards. For example, Switzerland is enforcing 19 percent capital ratios for its largest banks. These changes have boosted bank capital ratios throughout the advanced economies (figure 1).

Against this background, a key question for bank regulation is whether these reforms have gone too far or not far enough. Put differently, what is the socially optimal bank capitalization? Providing an answer to this question in its general form is likely impossible. That would require defining a social welfare function, estimating the effect of bank capital on the cost and availability of credit, the probability and severity of banking crises, and the effect of banking crises on output and output volatility. The results of such an exercise would be highly dependent on the models and parameters chosen: the rigor would conceal a large degree of judgment.

This paper takes a different, less ambitious path to evaluating bank capital ratios. We ask what capital buffers would have been sufficient to absorb bank losses through equity in a large majority of past banking crises. Once the sufficient bank capital ratios are established, we verify that the costs of implementing such capital ratios appear acceptable. We call this "a seawall approach": akin to the analysis that goes into building a seawall, we assess the height of a wall that can absorb plausible large waves, then confirm that the cost of such a seawall is bearable. This reduced-form perspective has the benefit of reducing the number of assumptions relative to general equilibrium models, thus increasing the robustness and transparency of our findings.

Our seawall approach leads to robust results. We find that bank capital in the range of 15–23 percent of risk-weighted assets would have absorbed bank losses in most past banking crises in advanced economies. The marginal benefits of bank capital decline once banks' capital ratios reach 15–23 percent of risk-weighted assets. The reason is that the more extreme crises are rare and would require substantially more capital to manage them. The 15–23 percent inflection region for the marginal benefits of bank capital is robust to various alternative estimates—among others, using data on bank loan losses or on public recapitalizations of banks, considering bank loan losses or also securities losses, using industry-average bank capital or individual bank capital, etc. As such, a sharp inflection in the marginal benefits of bank capital in the 15–23 percent risk-weighted capital ratio range suggests that, for a wide range of possible costs of bank capital, optimal bank capital is within that range.

The inflection point in the marginal benefits of bank capital is less pronounced and occurs at higher capital ratios for emerging market and developing economies, where banking crises have often been associated with larger bank losses. This asymmetry highlights the complementarity of capital and institutional improvements (in regulation, supervision, and resolution) in reducing expected losses in a possible banking crisis. Further, reminiscent of the debate on sovereign debt sustainability, it stresses the correlation between the magnitude and frequency of macroeconomic shocks and the size of the buffers necessary to confront them.

It is important to be careful in relating our findings to capital regulation. First, our results are in terms of actual bank capital rather than minimum capital requirements. Banks tend to maintain buffers over minimum capital requirements, and can draw on those buffers in stressed periods. Second, tighter liquidity regulation and more stringent overall supervision (including detailed stress-testing) of banks in the wake of the crisis could have reduced bank risk for any given level of bank capital, compared with pre-crisis levels. Third, while this paper focuses exclusively on bank capital as a means to absorb bank losses, other bail-in-able instruments can contribute to loss-absorption capacity. For these reasons, optimal capital requirements can be somewhat below the 15–23 percent range identified in our analysis. This makes our findings consistent with the upper range of Basel III capital requirements and with the total loss-absorbing capacity (TLAC) standards for systemic bank institutions (Bank for International Settlements 2016).[1] Finally, we abstract from the beneficial effect higher capital might have on bank risk taking and expected nonperforming loans (NPLs).

The paper proceeds as follows. Section 2 reviews the literature on the benefits and costs of bank capital. Section 3 presents estimates of bank capital ratios sufficient to absorb losses in past banking crises using alternative "seawall" approaches. Section 4 discusses the robustness of these estimates. Section 5 reviews evidence on the costs of bank capital. Section 6 concludes.

2. Theoretical Background: The Benefits and Costs of Bank Capital

Higher bank capital has several benefits from a financial stability perspective, but might also impose costs on banks and society. In an idealized Modigliani-Miller (1958; henceforth MM) world without tax deductibility of interest rate costs, bankruptcy costs, or agency problems, bank leverage does not affect social welfare (or bank profits). In this world, capital requirements are at the same time costless and irrelevant. In practice, however, several frictions imply that the MM paradigm does not apply (at least to banks), and that capital may affect the way banks behave and their profitability. In particular, asymmetric information entails significant agency problems and externalities magnify the social cost of bank failure. Then, capital

[1]See Zhou et al. 2012 and Flannery 2014 for a discussion of the relative benefits and costs of bail-in-able instruments compared with bank equity in providing loss-absorption capacity.

can play an important role in aligning banks' incentives with social welfare.

2.1 Benefits

First, capital serves as a buffer that absorbs losses and reduces the probability of bank failure. This protects bank creditors and, in systems with explicit or implicit public guarantees, taxpayers. Second, capital has a preventative role by improving incentives for better risk management. When asymmetric information prevents creditors from pricing bank risk-taking at the margin, banks operating under the protection of limited liability will tend to take excessive risks. Capital can limit these excesses by increasing shareholders' "skin in the game"—the amount of equity at risk in the event of bank failure (Marcus 1984; Myers and Majluf 1984; Keeley 1990; Esty 1998; Hellman, Murdock, and Stiglitz 2000; Matutes and Vives 2000; Repullo 2004). This includes the role of bank capital in helping minimize market discipline distortions associated with deposit insurance and implicit government "too-big-to-fail" guarantees.[2]

Market forces push banks to maintain some positive level of capital. For example, higher capital helps banks attract funds (Holmstrom and Tirole 1997), maintain long-term customer relationships (Allen, Carletti, and Marquez 2011), and carry risks essential to lending (Calem and Rob 1999; Perotti, Ratnovski, and Vlahu 2011). However, it is widely accepted that these forces are not sufficient to ensure that the market equilibrium bank capital levels deliver a welfare-maximizing allocation. Put differently, the frictions discussed above imply that the private return to capital is lower than the social return. Thus, banks will tend to hold less capital than what is socially optimal. This provides a rationale for regulation aimed at increasing bank capital relative to the laissez-faire equilibrium (this typically comes in the form of risk-weighted minimum capital requirements and more recently of caps on leverage ratios).

[2]Some have argued that how bank ownership is distributed matters. When bank equity is held by outside investors with high risk preferences or concentrated ownership, it may increase bank risk-taking (Laeven and Levine 2009). Indeed, passive outside investors are likely to have a limited disciplining role.

2.2 Costs

In analyzing the costs of bank capital, it is important to distinguish between the transition and steady-state impact of higher capital requirements. The costs associated with the transition to heightened capital requirements are not relevant at the steady state. These are costs stemming from raising new external equity or/and reducing the growth of assets. Equity issuance is subject to non-negligible underwriting fees, usually of 5–7 percent. Also, there are signaling costs: issuing equity may require substantial discounts when incumbent investors and managers have information about the firm that new equity investors do not have (Myers and Majluf 1984). Therefore, one would expect that any rapid increase in mandatory capital ratios would take place at least partially through an adjustment of bank assets, with potentially large negative effects on credit and macroeconomic performance.

In principle, the signaling cost should be low for gradual, regulatory increases in bank capital that affect the whole financial system (and therefore imply no hidden information on individual bank conditions). And the transition costs could be mitigated by giving banks time to adjust their balance sheets gradually. This might enable banks to increase capital using retained earnings or external capital issuance timed to beneficial market conditions. The caveat is that the benefits of regulatory gradualism might be limited if market pressures force banks to adjust to new capital standards rapidly.

The steady-state costs of higher capital requirements are those that occur after a permanent change in the funding mix of banks is completed. Some of the costs associated with a heavier reliance on equity are similar for banks and nonfinancial firms. For example, in many jurisdictions, debt has a more favorable tax treatment then equity (De Mooji 2011). Aside from tax issues, equity can be costlier if, due to some frictions, a decrease in leverage does not lower the banks' required return on equity. Yet the nature of these frictions requires a deeper discussion to understand the social welfare implications of higher bank capital.

The most notable reason why lower bank leverage may not pass through into a lower required return on equity is that deposits and other debt liabilities often benefit from subsidized safety net protections, including deposit insurance and too-big-to-fail subsidies that benefit bank debt more than bank equity (Kane 1989). Junior debt holders and uninsured depositors suffered minimal losses during the recent crisis, especially when compared with shareholders. As a result, banks' overall costs of funding may increase with greater equity finance. Yet this increase in banks' cost of funding is primarily a private cost to banks. While it might affect the cost and availability of bank credit, at the same time it reduces the distortions associated, for example, with the fiscal and incentives effects of expected bailouts.

Other notable costs stem from the fact that whereas for a nonfinancial firm leverage is a funding decision, for a bank its debt is also an output. The literature suggests that some economic agents, so-called cash investors, value bank debt for its high (often immediate, for deposits) liquidity and safety. When banks replace debt with equity, this destroys some economic value intrinsic to bank debt (Song and Thakor 2007; DeAngelo and Stulz 2013; Allen, Carletti, and Marquez 2015).[3] This reduces the cash investors' surplus, along with bank profits, and can harm bank borrowers through a higher cost of credit.

The existing literature has put forward several reasons why some investors value liquid and nominally safe assets such as bank debt. The hypotheses include liquidity insurance and convenience (Bryant 1980; Diamond and Dybvig 1983; Gorton and Pennacchi 1990; Caballero and Krishnamurthy 2008), agency costs in the money management of corporations and sovereigns that make them eschew any investment risk (Caballero and Krishnamurthy 2008), or the usefulness of risk-insensitive claims as a transactions medium (Dang et al. 2017). Empirical studies document the demand for safe and liquid assets (Gorton, Lewellen, and Metrick 2012), confirming the presence of cash investors in financial markets.[4] Krishnamurthy and Vissing-Jorgensen (2012) and Greenwood, Hanson, and Stein (2015) estimate the risk-adjusted premium of Treasuries over other bonds

[3]Of course, banks need to maintain some equity to ensure that their debt is safe and liquid in most states of the world (Hellwig 2014). But equity above that would crowd out socially valuable bank debt.

[4]The hypothesis of investors' preference for safe and liquid assets is often used also in the analysis of shadow banking; see Gennaioli, Shleifer, and Vishny (2012, 2013) and Claessens, Ghosh, and Mihet (2013).

to be −50 to −70 basis points; this can be taken as an estimate of the funding cost advantage inherent in safe and liquid bank debt too.

Finally, a related but separate issue is the role short-term debt can play in disciplining banks (Calomiris and Kahn 1991; Diamond and Rajan 2000; Kashyap, Rajan, and Stein 2008). This relates more to the composition of bank debt than bank leverage per se. The argument is that without demandable debt that gives creditors the ability to "run" on weak banks, banks would engage in riskier behavior. However, the crisis has led some to question the role that short-term debt can play in protecting financial stability: it provided little discipline before the crisis but contributed to extreme, across-the-board, runs once the crisis hit (Krishnamurthy 2010; Huang and Ratnovski 2011; Gorton and Metrick 2012). Moreover, it is unclear why market discipline cannot be provided by only small amounts of short-term bank debt.

2.3 Systemic Implications

The analysis of the costs and benefits of bank capital acquires additional dimensions when the focus shifts from the stability of individual institutions to that of the financial system as a whole. Individual bank distress may propagate to other banks through direct interbank exposures, fire sales, and contagious panics (Allen and Gale 2000; Gale and Özgür 2005; Admati et al. 2010; Admati and Hellwig 2014). Then, a bank's higher capitalization, by reducing the probability of its distress, helps avoid the associated systemic spillovers.

Moreover, competitive pressures may act as a systemic multiplier of the beneficial effects of individual banks' capital. Weak or "zombie" banks taking excessive risks (including by reducing lending standards and intermediation margins) may force healthy ones to engage in similar practices to protect their market share. And, to the extent that bank shareholders and creditors cannot fully evaluate a bank's risk-adjusted performance, similar pressures will bear as bank managers at healthy banks attempt to match the riskier banks' profitability (Caballero, Hoshi, and Kashyap 2008).

The level and distribution of capital across a banking system may also matter. Sufficient aggregate capital may enable strong banks to acquire weak institutions (Acharya, Engle, and Richardson 2012). In other cases, healthy banks may curtail lending if they expect macroeconomic conditions to be negatively affected by the reduction in credit supply due to weakness at other banks (Bebchuk and Goldstein 2011). Related, the risk of contagion associated with weakness at a systemic bank may reduce the incentives for acting prudently at other banks (Dell'Ariccia and Ratnovski 2013).

On the cost side, higher bank capital requirements may affect the allocation of activities across different financial intermediaries. In particular, "too high" capital requirements may trigger a migration of activities from banks to less-regulated parts of the financial system and thus increase systemic risk (Goodhart 2010; Martin and Parigi 2013; and Plantin 2015).[5]

2.4 The Balance of the Benefits and Costs of Higher Bank Capital

The papers that examine the balance of the benefits and costs of higher bank capital reach varying conclusions, depending on their assumptions and methodologies. At one end of the spectrum, Admati et al. (2010) and Admati and Hellwig (2014) argue that higher bank capital is not socially costly thanks to a nearly complete MM offset. Moreover, they argue that higher bank capital enhances the liquidity of bank debt. This compensates for its lower volume and keeps the value of safe asset premiums in the banks' cost of funding essentially constant. Overall, they argue for a 20 percent bank leverage ratio, corresponding to a 35 percent bank risk-weighted capital ratio. Our analysis suggests that additional capital beyond the 15–23 percent range has low benefits in terms of preventing additional banking crises.

[5]Another argument against "too high" bank capital is that while higher capital requirements reduce bank risk, they may at the same time increase borrowers' risk. This may occur when high capital requirements dampen interbank competition, increasing the cost of credit and inducting risk-shifting by bank borrowers (Boyd and De Nicolo 2005; Hakenes and Schnabel 2011). Or it may occur when high capital makes banks tolerant of risky lending, and firms take risk without the fear of being denied credit (Gornall and Strebulayev 2018).

On the opposite end of the spectrum, the Institute of International Finance (IIF) (2010) argues that any increases in bank capital can substantially affect the availability and cost of bank credit. In this argument, the focus is heavily on the transitional costs of higher bank capital, highlighting two concerns. The first is that heavy equity issuance may be costly during times of high economic and regulatory uncertainty. Whereas the IIF conjectures that increases in bank capital are therefore undesirable, we suggest that the cost of equity is minimized when higher bank capital requirements are imposed gradually. The second IIF concern is that the past costs of equity issuance affect the required rate of return on bank equity going forward. This argument appears less valid when the increase in bank capital is a one-off event. Overall, as we highlight in our analysis, the relatively healthy post-crisis dynamics of bank credit do not validate the IIF (2010) views.

Among other papers, Miles, Yang, and Marcheggiano (2013), Firestone, Lorenc, and Ranish (2017), and Barth and Miller (2018) use a common method but alternative data sets to assess the benefits and costs of bank capital. The benefits of bank capital are estimated from historical data (including those from the 19th century) on the link between bank capital, the probability of banking crises, and output losses in banking crisis. The caveat is that historical data may misrepresent the contemporary link between bank capital and the probability of banking crises, while the estimation of output losses relies on the assumptions regarding the counterfactual. Both papers allow for an incomplete MM offset in the cost of bank capital. The estimated optimal risk-weighted bank capital is 16–20 percent in Miles, Yang, and Marcheggiano (2013), within our range; 13–26 percent in Firestone, Lorenc, and Ranish (2017), wider than our range; and 26 percent in Barth and Miller (2018), above our range. The Bank for International Settlements (BIS) (2010) and Cline (2016) use a common model-based method to assess the benefits of higher bank capital, and also allow for an incomplete MM offset in the cost of bank capital. Their estimates of optimal bank capital are 10–12.5 percent and 12–14 percent, respectively, below our range. The Federal Reserve Bank of Minneapolis (2017) replicates the analysis of our paper and suggests optimal bank capital of 23 percent, at the top of our range.

3. Bank Capital: A Seawall Approach

The analysis above suggests that bank capital levels position a banking system on a tradeoff between financial stability and the cost of financial intermediation. Implicit in this tradeoff is the notion that there exists an "optimal" level of capital that maximizes some aggregate welfare function with output growth and volatility as ultimate arguments and bank stability and the cost and availability of credit as intermediate ones.

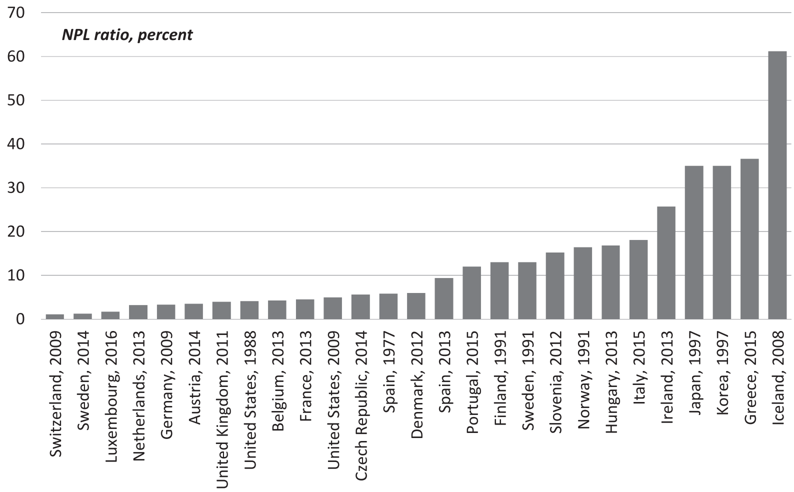
Estimating this optimal level of bank capital is, however, a complex task. It requires defining a social welfare function, estimating the effect of bank capital on the cost and availability of credit, the probability and severity of banking crises, and the effect of credit availability and banking crises on output and output volatility. Such an exercise would require simplifying assumptions, likely making its results too model-, bank-, and sample-specific to provide convincing policy guidance.

Against this background, we take a different, less ambitious path to evaluating the pros and cons of higher/lower bank capital ratios. We ask what capital buffers would have been sufficient to absorb all bank losses through equity in most of past banking crises. We think of this as a "seawall approach": choosing the height of a seawall that historically most often would have been sufficient. We perform this analysis based on two types of data. First, in section 3.1, we use data on nonperforming loan ratios in past banking crises and ask how much bank capital would have been enough to absorb the loan losses. Second, in section 3.2, we use data on public recapitalizations of banks and ask how much capital would have been enough to prevent those. It turns out that the results based on two different types of data are very similar, giving us comfort as to the robustness of our findings. Once the "sufficient" bank capital ratios are established, we verify in section 5 that the costs of implementing such capital ratios are acceptable.

3.1 Bank Capital Sufficient to Absorb Bank Loan Losses

The first approach considers the capacity of banks to absorb loan losses. We consider NPL ratios in past banking crises and ask how much capital banks would have needed to absorb loan losses and

Figure 2. Peak NPL Ratios in Banking Crises in OECD Countries



maintain positive equity, thus avoiding losses to creditors (cf. Ratnovski 2013). We base the analysis on the historic banking crises NPL data for OECD (Organisation for Economic Co-operation and Development) countries from a newly updated data set by Laeven and Valencia (2018). Figure 2 summarizes this data, showing peak NPL ratios during respective crises.[6]

To convert NPLs into bank capital ratios needed to absorb loan losses, we proceed in four steps. Each step is associated with assumptions on specific conversion parameters. We rely on the literature to establish the baseline parameters and the confidence range (stressed scenario) parameters. Overall, this approach implies that the resulting "sufficient" loss-absorbing capital is better discussed in terms of ranges rather than point estimates.

First, we convert NPLs into loan losses, by adjusting the nonperforming loan ratio for loss given default (LGD). Unfortunately, there is limited cross-country data on loss given default. For the baseline, we use estimates for the United States suggesting that the mean

[6]We abstract, in this exercise, from potential differences in accounting and prudential requirements regarding NPLs across countries and over time, an issue that is hard to circumvent.

loss given default over 1970–2003 was about 50 percent on average in normal times (Schuermann 2004; Shibut and Singer 2014; Johnston Ross and Shibut 2015). In a robustness scenario, we allow for LGD of up to 75 percent, to reflect possibly higher LGD in systemic banking crisis times. Second, we establish a share of bank losses that can be absorbed by prior provisioning. In the United States, loan loss provisioning averaged about 1.5 percent historically. In Spain, dynamic provisioning achieved similar buffers prior to the 2008 financial crisis (Saurina 2009). Therefore, loan loss reserves of about 1.5 percent seem a reasonable assumption. Third, we compute capital ratios that would enable banks to absorb the estimated losses and remain in positive equity. For this, we take bank capital equivalent to loan losses net of provisions and add an additional 1 percent of capital as a margin of safety in the baseline. The idea is that positive post-crisis bank capital enables the government to sell off the weak bank without a fiscal injection (which is essential to sell a bank with negative capital). In a robustness scenario, we allow for a higher margin of safety. Finally, we convert resulting unweighted capital needs into risk-weighted capital by applying a 1.75 ratio of total assets to risk-weighted assets (RWA), corresponding to the average such ratio for U.S. banks (Avramova and Le Lesle 2012). In another robustness scenario, we use a 2.5 ratio, corresponding to this ratio in Spanish banks (the highest in OECD countries).

Overall, the baseline formula that converts loan losses in a banking crisis into the risk-weighted capital ratios needed to absorb them is

$$\begin{aligned}\text{Bank capital} = {} & (\text{NPL} * \text{LGD} - \text{Provisions} + \text{Margin of safety}) \\ & * (\text{TA} \,/\, \text{RWA}).\end{aligned}$$

Table 2 illustrates our calculation. In the baseline (column 1), a hypothetical 18 percent NPL ratio corresponds to 9 percent loan losses, and loan losses net of provisions of 7.5 percent of total loans. To cover with a margin of safety 7.5 percent loan losses net of provisions, a bank needs an 8.5 percent leverage ratio, corresponding to approximately 15 percent risk-weighted capital ratio. Higher LGD, lower risk weights, or higher margins of safety (columns 2–4) increase the corresponding bank capital ratio to up to 23 percent.

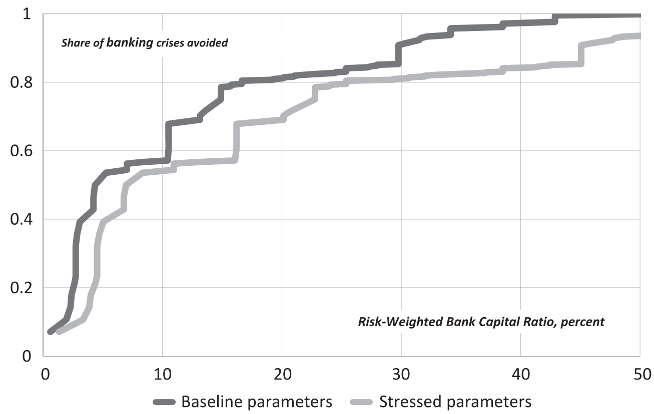
Table 2. Example: Capital Needed to Absorb NPLs Equal to 18 Percent of Assets (all values in percent)

	Baseline (1)	**Higher LGD (2)**	**Higher TA/RWA (3)**	**Higher M.O.S. (4)**
NPL during a Banking Crisis	18.0	18.0	18.0	18.0
Loss Given Default	50.0	**75.0**	50.0	50.0
Loan Losses (1*2) (Mean Point)	91.0	13.5	9.0	9.0
Absorbed by Prior Provisioning	1.5	1.5	1.5	1.5
Loan Losses Net of Provisions (3–4)	7.5	12.0	7.5	7.5
Margin of Safety (Residual Capital)	1.0	1.0	1.0	**3.0**
Capital-to-Assets Ratio or Leverage Ratio (5+6)	8.5	13.0	8.5	10.5
Total Assets/RWA	175.0	175.0	**250.0**	175.0
Capital Ratio (Percent of RWA) (7*8)	14.9	22.8	21.3	18.4

We then use the transformation above applied to the distribution of NPL ratios in past banking crises to transform a given bank capital ratio into the share of banking crises in which it could fully absorb loan losses. Figure 3 reports the share of advanced-economy banking crises in which banks would have maintained positive equity as a function of hypothetical bank risk-weighted capital ratios. We plot the function for the baseline (light gray line) as well as for a stressed scenario corresponding to LGD of 75 percent (as in column 2 of table 2, darker line). A line for a stressed scenario corresponding to lower risk weights (column 3 of table 2) is similar to the higher-LGD scenario.[7] A higher margin of safety (column 4 of table 2) yields results

[7]To the extent that such higher total assets to risk-weighted assets conversion ratio corresponds to safer bank portfolios, it would arguably correspond to lower loss-given-default estimates, thus partly compensating for its direct effect on bank capital needs. Note that we use historic conversion ratios that account

Figure 3. Share of Banking Crises without Creditor Losses in OECD Countries



in between the baseline and the stressed scenario; varying the margin of safety within reasonable bounds does not significantly affect the range of estimated bank capital needs.

The baseline schedule suggests that, in OECD countries, the marginal benefit of additional capital from a loss-absorption point of view is relatively high until a 15 percent risk-weighted capital ratio (which enables banks to absorb losses in more than 80 percent of banking crises). The marginal benefit of additional capital declines rapidly after that. This means that attempting to absorb bank losses in a few exceptionally extreme crises requires very high capital ratios. For the stressed scenario, the inflection point is close to 23 percent of risk-weighted assets. Notably, the baseline result is relatively close to the model-based estimates in BIS (2010) that suggest that capital of 15 percent would avoid imposing losses on creditors in about 90 percent of banking crises.[8]

for the fact that some banks may engage in "strategic risk-weighting" (Enrich and Colchester 2012; Mariathasan and Merrouche 2014).

[8]It is useful to compare the parameters underlying our analysis with the parameters from recent U.S. and European bank stress tests. In the 2015 U.S. stress test (Board of Governors of the Federal Reserve System 2015), the "severely adverse scenario" led to 4.5 percent loan losses and a loss of 5.5 percentage points of risk-weighted bank capital. In the 2014 European Banking Authority (EBA) stress test (EBA 2014), the "adverse scenario" led to 2.3 percent loan losses that

3.2 Capital Sufficient to Avoid Public Recapitalizations of Banks

The next approach uses alternative data and considers how much capital banks would have needed to avoid public recapitalizations during past crises. The working assumption is that, historically, post-crisis bank recapitalizations brought banks to the minimum level of capital needed for viability. If this assumption is correct and if, prior to the crisis, banks had had capital equivalent to the sum of actual pre-crisis capital and the post-crisis public capital injection, then, other things remaining equal, no public recapitalizations would have been required.

We combine data from Bankscope on average capital ratios by country in 2007 in banking systems of OECD countries that experienced a crisis over the period 2007–13 with data from Laeven and Valencia (2018) on the fiscal outlays associated with bank recapitalizations (with both variables being expressed as percentages of total risk-weighted assets of the banking system in each country). The sum of pre-crisis bank capital levels and public bank recapitalization injections is shown in figure 4.

Figure 5 uses the data summarized in figure 6 to relate hypothetical pre-crisis bank capital levels to the share of public recapitalization events that they would have helped avoid. Consistent with our previous findings, the marginal benefit of additional capital in terms of avoiding public recapitalization episodes is relatively high until 15–17 percent risk-weighted capital ratios (which help avoid public recapitalizations in 75 percent of banking crises). The marginal benefit of bank capital declines rapidly after that. The results from data on public bank recapitalization expenses are therefore highly consistent with our previous estimates based on NPL data—both confirm

induced a loss of 4.4 percentage points of bank capital. Two observations are worth emphasizing: First, loan losses corresponding to the 85th percentage of banking crises (as used in our analysis) are two to four times as high as losses in the U.S. and European stress tests (which, being stress tests over plausible outcomes, consider relatively milder scenarios than those of a full-blown crisis). This highlights the conservative nature of our estimates. Second, our analysis employs a coefficient of 1.75 to convert loan losses into losses in risk-weighted capital; this is in between the coefficients of 1.25 and 1.9 implied by U.S. and European stress tests. (A smaller coefficient in U.S. stress tests is a product of low predicted losses on bank securities holdings.)

Figure 4. Pre-crises Bank Capital and Fiscal Recapitalization Expenses in Banking Crises from 2007 Onward

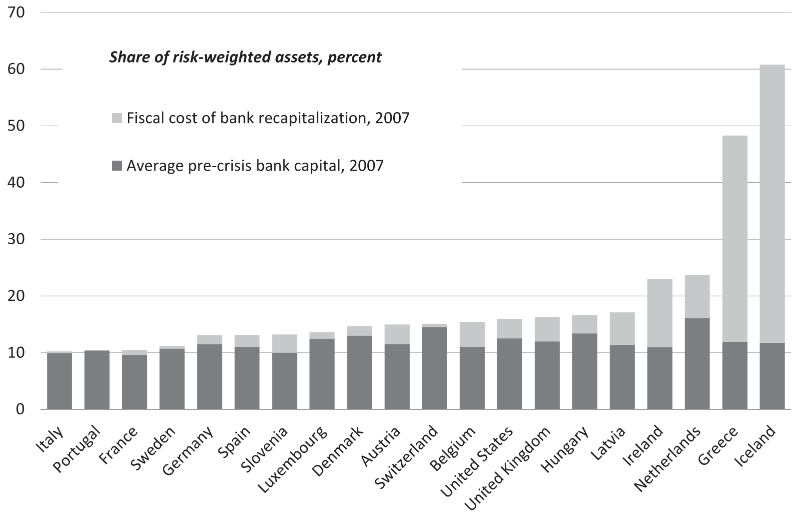
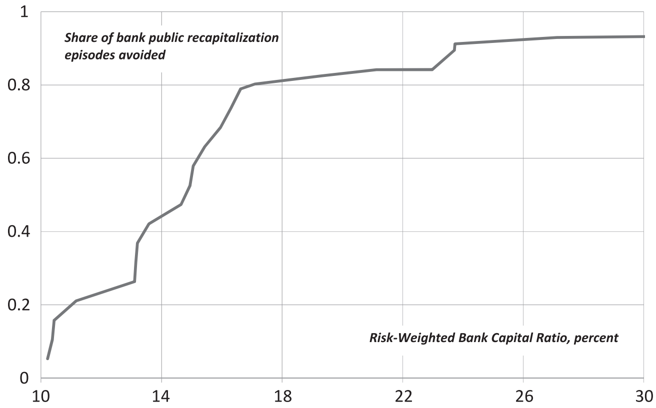


Figure 5. Share of Public Recapitalization Avoided, Depending on Hypothetical Pre-crisis Bank Capital Ratios



that bank capital in the range of 15–23 percent of risk-weighted assets would have enabled banks to absorb loan losses and would have allowed avoidance of public recapitalizations of banks in about 80 percent of banking crises in advanced economies.

4. Robustness

The previous section estimated bank capital levels that would have been sufficient to absorb losses in most historic banking crises in advanced economies. We offered estimates based on two alternative types of data: NPLs and public bank recapitalizations. The fact that the alternative methods produced very similar results helps assuage concerns related to the parameter uncertainty in our analysis. In this section, we explore the robustness of our results further.

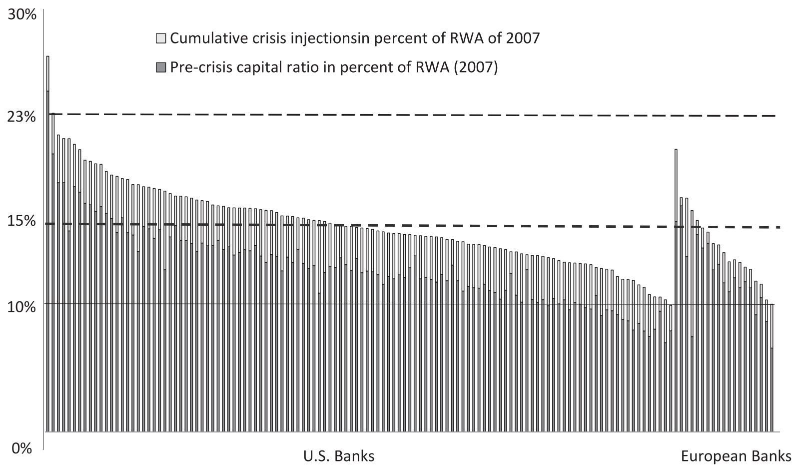
4.1 Loan Losses versus Securities Losses

Our analysis based on NPL ratios equates bank losses with loan losses; more precisely, we assume that banks accrue losses uniformly across different assets on their balance sheet. This implies that we will tend to overestimate capital needs when losses are concentrated on loans and underestimate them when they are concentrated on securities (such as was the case for many advanced economies in the global financial crisis). However, securities losses during the global financial crisis (Berrospide 2013) and in the Federal Reserve's "severely adverse scenario" stress tests (Board of Governors of the Federal Reserve System 2015) are close to or below the loan losses, suggesting that loan loss estimates can be extrapolated to securities losses. Further, both loan losses and securities losses during the global financial crisis and in the "severely adverse" stress test are below our estimates of the 80th percentile of loan losses in OECD banking crises, suggesting that our estimates of the capital needs are in any case sufficiently conservative.

4.2 Aggregate Losses versus Losses in Individual Banks

Our analysis is based on average NPL ratios and average bank recapitalization expenditures in the banking system. In practice, losses or

Figure 6. Pre-crisis Bank Capital and Capital Injections during the Crisis

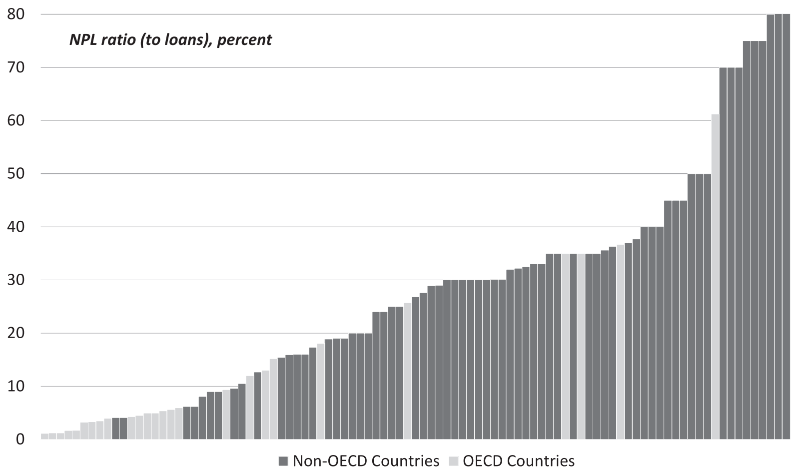


recapitalization needs at individual banks may differ from the country mean, and thus our methodology will underestimate the level of capital necessary to preserve positive equity across the entire system. (Still, the average level of capital remains informative as to the overall capacity of the banking system to absorb losses and may affect the authorities' ability to confront the crisis, including by facilitating takeovers of weaker banks by stronger ones.) To assess the robustness of our findings to the distribution of bank capital needs, we examine bank-level government capital injections during the recent crisis in large European and U.S. banks.[9]

Figure 6 plots, at the bank level, the sum of the pre-crisis capital and capital injections during the crisis (both in percent of pre-crisis RWA, similar to figure 4). The figure suggests that a capital of 15 percent in 2007 would have avoided the need for capital injection in almost 55 percent of cases in the United States and 75 percent

[9]The data on capital injections in European banks are taken from estimates by Fratianni and Marchionne (2013), merged with bank financials from SNL Financial, and only cover injections between November of 2008 and January of 2010. The data on U.S. injections are from SNL Financial and are based on the Troubled Asset Relief Program.

Figure 7. Nonperforming Loans as a Share of Loans: OECD Countries versus Non-OECD Countries

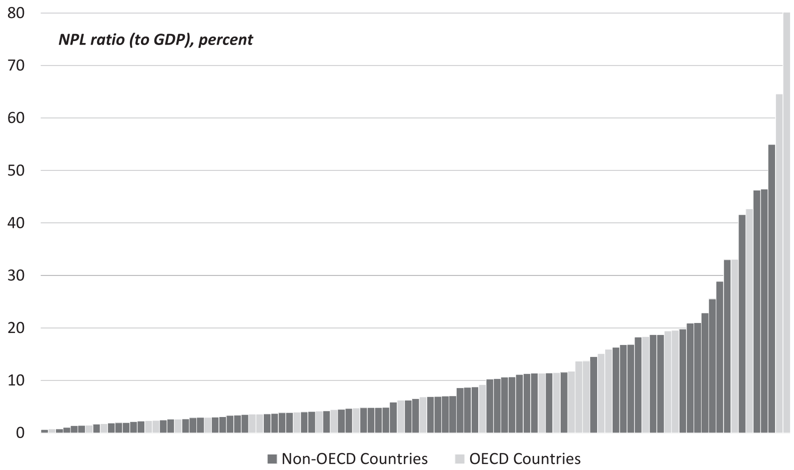


of cases in Europe (based on the available data), while a capital of 23 percent would have eliminated the need for injection in virtually all cases. While the 55 percent figure in the case of the United States might seem low, note that this is based on the lower bound of our range. Further, the Capital Purchase Program's terms were relatively attractive to avoid stigmatizing participating banks as being weak (Swagel 2009).

4.3 Advanced Economies versus Emerging Markets and Developing Countries

The analysis of section 3 focused on advanced economies only for comparability. Emerging market and developing countries have historically had higher NPL ratios in banking crises (figure 7, light gray bars). In principle, higher NPLs, all else being equal, call for higher levels of capital to absorb them. Applying the method of section 3.1 to convert NPLs into bank capital needed to absorb them suggests that capital ratios in the 15–23 percent range would have been sufficient to absorb losses in only about half of all banking crises in non-OECD economies. This is not surprising once one observes that in these countries, macroeconomic shocks tend to be larger, credit

Figure 8. Nonperforming Loans as a Share of GDP: OECD Countries versus Non-OECD Countries



tends to be less diversified, and institutional factors lead to larger loss-given-default ratios.

Consistent with this view, non-OECD countries (on average) have been imposing higher capital requirements on banks. In 2010, the minimum capital ratios in OECD countries were almost uniformly 8 percent. In contrast, the median minimum capital ratio in non-OECD countries was 10 percent. Moreover, almost a quarter of non-OECD countries had a minimum capital ratio of 12–15 percent, or 50 percent higher than what was typical in OECD countries (Barth, Caprio, and Levine 2013).

An offsetting factor to higher bank losses during banking crises is that, historically, non-OECD countries tended to have much smaller banking systems relative to gross domestic product (GDP) than OECD countries. This leads to NPLs as a share of GDP rather than as a share of bank loans being comparable in banking crises in advanced and other economies (figure 8). This means that when bank losses exceed the absorption capacity provided by capital, their effect on the economy (and thus the fiscal accounts) is likely also to be smaller. Everything else being equal, the ex post cleanup operations are likely to be less onerous than in countries with larger banking systems. For instance, given the smaller size of their banking systems, had non-OECD countries imposed bank capital ratios in the 15–23 percent range, in 80 percent of banking crises losses exceeding the absorption capacity of bank capital would have been within 3 percent of GDP, usually a manageable fiscal burden. Consequently, desirable bank capital levels in OECD and non-OECD countries might be closer than they appear from the first-brush NPL ratios analysis.

There are, however, two caveats to this conclusion. First, the potential fiscal costs of bank cleanups in emerging economies are destined to increase beyond this estimate, as many have recently experienced rapid credit growth, making current ratios of bank credit to GDP higher than past averages suggest. Second, the estimate should not give rise to complacency, as bank losses in the remaining 20 percent of banking crises (often, twin crises) would have been substantial. Further, if we took the higher capital needs in non-OECD countries at face value, a strategy complementing higher capital ratios would be to reduce potential NPLs through institutional improvements (in regulation, supervision, and resolution).

4.4 Dealing with Severe Banking Crises

In focusing on the banks' ability to absorb losses in most banking crises, our approach omits the issue of cleanup costs in the left tail of most severe banking crises where even higher bank capital does not provide sufficient loss-absorption capacity. While rare, the severe crises might also be the costliest. In a way, this omission is an inherent limitation of our seawall approach: a real seawall also might not protect from a truly large tsunami or an earthquake. Ex post government intervention might be essential in truly severe crises (Geithner 2014). Still, even for very severe banking crises, high bank capital would decrease ex post cleanup costs and provide the authorities additional breathing space for a more orderly resolution.

5. The Costs of Higher Bank Capital

Section 3 showed that capital in the range of 15–23 percent of risk-weighted assets would have absorbed bank losses in the majority of past banking crises, and higher capital ratios are relatively ineffective in preventing additional banking crises. The inflection of marginal benefits of bank capital points to capital in the 15–23 percent range as a plausible "optimal" bank capital that may enhance financial stability and reduce expected crisis-associated fiscal outlays. Yet, we said nothing about the effects that this higher capital would have on the availability and cost of bank credit, and, ultimately, on macroeconomic performance. In this section, we verify that the costs of bank capital in the 15–23 percent range appear manageable. To this end, we review the literature on the costs of bank capital (summarized in table A.1 in the appendix), and discuss in this section its main findings.

5.1 Steady-State Cost of Capital

It is useful to separate the steady-state and the transitional costs of higher bank capital. The econometric literature on the steady-state costs of bank capital is relatively thin, reflecting the difficulty in estimating such costs. Because of relatively stable capital regulation over the past few decades, most econometric studies exploit the cross-sectional and time-series variation of bank capital within a given regulatory framework. Here, based on U.S. data, the literature generally finds that a 1 percentage point higher tier 1 capital ratio is associated with loan rates that are 2.5 basis points higher: a very modest relationship. A caveat is that, since this variation reflects banks' endogenous choices (banks optimizing how much capital to hold over the regulatory minimums), we can expect the effect of an exogenous, regulatory-mandated increase in capital to be larger than the econometric estimates obtained in such studies.[10]

Given the limitations of econometric frameworks in the absence of exogenous variation in bank capital, other studies rely on calibrated models to assess the steady-state costs of bank capital. The key parameter of these models is the degree of the assumed Modigliani-Miller offset: to what extent a policy-imposed increase in the capital requirement would increase the total funding costs of banks. An increase in a firm's capital reduces its riskiness and thus

[10]These estimates are consistent with prima facie historical evidence, showing that spreads between the reference and lending rates were not higher in periods when banks were much more highly capitalized (see Miles, Yang, and Marcheggiano 2013 for evidence from the United Kingdom and the United States).

its cost of borrowing. As discussed in section 2, Modigliani and Miller (1958) show that under a set of ideal assumptions, this effect fully offsets any potential increase in the total funding cost from a shift in funding structure, making an increase in capital essentially costless to banks and inconsequential to lending rates. Therefore, the degree to which MM holds is a crucial question. Overall, most studies that allow for some MM offset find extremely small effects: the impact of a 1 percentage point increase in capital requirements on lending rates is about 2 basis points (bps), consistent with empirical studies. Studies that assume that MM offset does not exist suggest an effect of up to 13 bps—still a reasonably modest one. The effects of this increase in the cost of bank funding on lending can be compensated in the long term, for example, through more accommodative monetary policy.

The effect of higher bank capital on safe assets creation also appears modest. Assume that bank balance sheet size is fixed, and the increase in bank capital reduces one-for-one the volume of bank deposits. Then, an increase in bank capital from 14 percent of risk-weighted assets, an average across large banks today, to 23 percent, the top of our range, implies a reduction in the volume of bank deposits by (23–14)/1.75 = 5.1 percent of bank liabilities (where 1.75 is the average risk weight; see table 2). Forgoing a 70 bps safety premium possibly associated with bank deposits (set at the top of the range in Krishnamurthy and Vissing-Jorgensen 2012 and Greenwood, Hanson, and Stein 2015) on that volume of bank liabilities would increase bank cost of funding by a modest 3 bps. This brings the total effect of higher bank capital on the cost of lending to 5–16 bps, depending on the MM offset.[11]

Modest steady-state costs of higher bank capital imply low additional incentives to migrate activities out of the regulated banking system. This mitigates the theoretical concerns about a possible expansion of shadow banking. Indeed, there is little evidence

[11]Some argue that the economy exhibits a relatively inelastic demand for safe assets (Gorton, Lewellen, and Metrick 2012; Gorton and Ordonez 2013). This raises the question of what assets can substitute for a reduced volume of bank deposits in safe assets supply. Yet this concern also appears manageable. For example, in the United States, a decline in safe assets supply of the magnitude of 5.1 percent of bank liabilities can be compensated by increasing the volume of outstanding Treasury bills and notes by 8 percent (all data as of end-2017).

of migration to shadow banking in countries implementing higher than Basel leverage requirements, such as Switzerland. Further, to the extent that the migration of activities to shadow banking represents regulatory arbitrage, it can be reduced or prevented by more comprehensive bank regulation (Claessens and Ratnovski 2014).

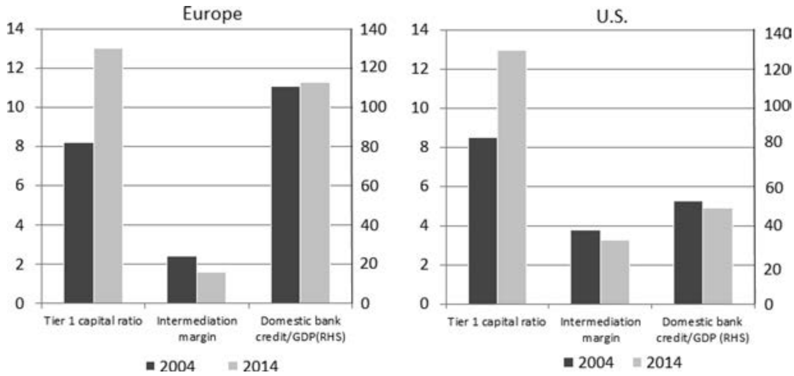
Based on these estimates, it would be relatively easy to argue for even higher bank capital ratios than the 15–23 percent range suggested in the previous section. Still, given that banking crises of extreme magnitudes are rare, it is difficult to see such extremely high levels of bank capital as a policy priority. Also, there is some risk that the costs of bank capital may increase nonlinearly in the level of capital—for example, for some adverse effects from its impact on bank business models. This calls for caution in extrapolating observed low costs of capital into the costs of capital at substantially higher levels.

5.2 Costs of Transitioning to Higher Capital

A larger empirical literature (generally employing tighter identification strategies than the literature on steady-state costs) documents the transitional costs of changes in bank capital. This literature often exploits bank-level shocks to capital—resulting either from losses or idiosyncratic (bank-level) regulatory actions—to identify the exogenous effects of tighter capital regulation on the availability and cost of bank credit. The literature finds that a 1 percentage point increased capital requirement is associated with a 5–8 percentage point contraction in lending volumes over the short run (see, for instance, Peek and Rosengren 2000; Aiyar, Calomiris, and Wieladek 2014; Brun, Fraisse, and Thesmar 2014; Eber and Minoiu 2015).

The problem with these estimates is that they rely on sudden changes in bank capital: events that mostly characterize banks that are in some state of distress. Many of the challenges associated with raising capital under these circumstances are not relevant for evaluating the effects of gradual changes in capital regulation that would affect an entire banking system. For instance, in the short run, distressed banks may be more likely to meet tighter regulatory requirements by reducing the asset side of their portfolios more than they would if they were fully sound and could raise capital gradually over time. Similarly, the stigma attached to a bank trying to raise capital

Figure 9. Bank Capital Ratios and Credit Provision, Pre- and Post-Crisis



in isolation is unlikely to apply in a context of systemwide regulatory reform. It follows that estimates based on short-lived bank-level shocks are likely to overestimate the transition costs of higher capital requirements. The literature also does not provide a guide as to how these transition costs vary depending on macroeconomic conditions and between rapidly growing emerging markets and advanced economies.

Also, there is evidence that transition costs tend to be lower if one allows banks to adjust to the new regime more gradually. For instance, calibrated models for several OECD countries suggest that, on average, over eight years, a transition to a 1 percentage point higher capital requirement is associated with a 17 bps increase in lending, a 1.5 percent decline in lending volume, and a 0.16 percent drop in GDP compared with the baseline (Macroeconomic Assessment Group 2010).

Consistent with these considerations, an analysis of the increase in capital requirements in the wake of the global financial crisis suggests that the effects of tighter regulation on intermediation margins and the overall supply of bank credit have been limited (Cecchetti 2014). For instance, average risk-weighted capital ratios at large banks in the United States and Europe increased by almost 5 percentage points between 2004 and 2014. But credit-to-GDP ratios and intermediation margins remained virtually unchanged (figure 9). Overall, this suggests that transition costs are likely manageable.

6. Conclusions

This paper contributed to the debate on the optimal capital levels in banks. We took a seawall approach. First, we considered how much capital would have been enough to absorb bank losses in a majority of historical banking crises. Second, we reviewed existing literature that suggests that the costs of such capital levels would be acceptable.

The key result of the paper is that bank capital in the range of 15–23 percent of risk-weighted assets would have absorbed bank losses in most past banking crises in advanced economies. This range was obtained based on two alternative sets of bank data (NPLs and public recapitalization), offering comfort that the results are robust to the modeling assumptions employed. Increases in capacity to absorb losses beyond this are likely to provide limited benefits. Hence, given the uncertainty surrounding the long-term welfare costs of bank capital, bank capital in the 15–23 percent range appears appropriate for banks in advanced economies.

The 15–23 percent estimate can be taken with a conservative bend. The estimate does not consider the potential reduction in risk-taking induced by higher capital through "skin-in-the-game" effects: the impact of higher bank capital on the incentives of bank shareholders and managers (Laeven and Ratnovski 2014). Further, capital requirements consistent with 15–23 percent bank capital may be below this range, as banks tend to hold buffers over the regulatory minimums, tighter post-crisis banking regulation in aspects other than bank capital (such as more stringent stress tests) may have reduced bank risk for any given level of bank capital, and because part of bail-in-able capacity might be provided by instruments other than bank capital (such as subordinated debt).

The 15–23 percent range is close to, if slightly above, the upper limit of the Basel III capital requirements, and is very similar to the Financial Stability Board's total loss-absorbing capacity standards, as well as the Federal Reserve's proposal of 9.5 percent of total leverage exposure for global systemic banks.[12] The range is also consistent with the 8 to 20 percent optimal bank capital range established in most calibrated dynamic general equilibrium models

[12] See http://www.federalreserve.gov/newsevents/press/bcreg/20151030a.htm.

(Van den Heuvel 2008; Nguyen 2013; Begenau 2014; Martinez-Miera and Suarez 2014; Mendicino et al. 2015).[13]

The estimated loss-absorption needs can be refined to allow for heterogeneity across banks and over time. Banks that are not systemically important (those that can be allowed to fail without major spillover effects) could be allowed to hold lower capital and loss-absorption capacity. Similarly, most banking crises follow periods of rapid credit growth, suggesting a role for countercyclical buffers (Borio 2014; Claessens 2015).

We find that the results on optimal bank capital are more nuanced for emerging markets and low-income countries. On the one hand, banking crises in these countries have historically been associated with greater bank losses. On the other hand, because banking systems in these countries tend to be smaller than those in advanced economies, losses in excess of capital will likely represent a smaller share of GDP and thus might have more limited macroeconomic effects. In this context, the relative role of greater loss-absorption capacity and improvement in governance and institutions aimed at reducing losses in crises should be the subject of future research.

In emerging market and developing economies, tighter bank capital standards can be complemented by institutional improvements (in regulation, supervision, resolution, and governance) to reduce possible losses in banking crises. In advanced economies, higher bank capital requirements may provide stronger incentives for regulatory arbitrage and increase the risk of activities migrating to unregulated or less regulated financial intermediaries (such as insurance companies or broker-dealers). In that context, it is essential that tighter capital and loss-absorption requirements are complemented with measures that widen the perimeter of prudential and macroprudential regulation.

The paper also reviewed empirical evidence on the costs of higher bank capital. The evidence overwhelmingly suggests that the steady-state (long-run) social costs of higher bank capital requirements,

[13]The calibrated welfare effects of varying bank capital within that range in terms of higher lending rates and lost consumption are relatively small. The relatively wide range of the estimates for the optimal level of capital reflects the many degrees of freedom in model design and parameter calibration. It is also consistent with a similarly wide 15–23 percent range of estimates obtained in our analysis.

within our estimated range, are likely to be small. The costs of transitioning to higher bank capital might be substantial, but also might be lower when capital adjustment is staggered or takes place in the upswing of the credit cycle. Overall, the cost of higher bank capital corresponding to our estimates appears acceptable.

Appendix

Table A.1. Estimates of the Steady-State and Transitional Effects of Higher Capital Requirements on the Cost of Bank Credit

Paper	**Data and Method**	**Cost of Capital**
	A. Steady-State Impact	
Elliott (2009)	U.S. banks, 2009:Q1–2009:Q2, calibration	4 pp ↑ in ratio of equity over unweighted assets leads to ↑ in lending rate by 80 bps.
Bank of England (2010)	U.K. banks, calibration.	1 pp ↑ in capital req. leads to ↑ in lending rate by 7 bps.
Bank for International Settlements (2010)	OECD banks, 1993–2007. Dynamic general equilibrium model.	1 pp ↑ in capital req. leads to ↑ in lending rate by 13 bps.
Berrospide and Edge (2010)	U.S. banks, 1992–2008. Empirical estimation.	1 pp ↑ in capital req. leads to ↑ in loan growth by 0.7–1.2 pp.
Kashyap, Stein, and Hanson (2010)	U.S. banks, 1976–2008. Calibration assuming tax effect and MM effect.	10 pp ↑ in capital req. leads to ↑ in lending rate by 25–45 bps.
King (2010)	OECD banks, 1993–2007. Calibration assuming no changes in return on equity and cost of debt and a full pass-through of cost of funding.	1 pp ↑ in capital req. leads to ↑ in lending rate by 15 bps.
de Resende, Dib, and Perevalov (2010)	Canadian banks. Dynamic general equilibrium model.	6 pp ↑ in capital req. leads to ↑ in lending rate by 7.5 bps, ↓ in lending by 0.24%, and ↓ in GDP by 0.07%.

(*continued*)

Table A.1. (Continued)

Paper	Data and Method	Cost of Capital
A. Steady-State Impact		
Santos and Winton (2010)	U.S. banks, 1987–2007. Empirical estimation. They test the hypothesis that less-capitalized banks charge higher rates on borrowers with no access to public debt markets.	5 pp ↑ in capital req. leads to ↓ in lending rate to credit-constrained firms by 15 bps.
Cosimano and Hakura (2011)	Banks in advanced economies, 2001–09. Empirical estimation.	1.3 pp ↑ in leverage ratio leads to ↑ in lending rates by 16 bps and ↓ in loan growth by 1.3% in the long run.
Schanz et al. (2011)	U.K. banks, 2006–09. Calibration.	1 pp ↑ in capital req. leads to ↑ in lending rate by 7.4 bps and ↓ in permanent GDP by 0.04%.
Slovik and Cournède (2011)	OECD banks, 2004–06. Calibration assuming no MM effect and a full pass-through.	1 pp ↑ in capital req. leads to ↑ in lending rate by 16 bps.
de Ramon et al. (2012)	U.K. banks, 1992–2010. Empirical estimation.	1 pp ↑ in aggregate capital ratio leads to ↑ in lending spreads by 9.4 bps.
Francis and Osborne (2012)	U.K. banks, 1996–2007. Empirical estimation.	3 pp ↑ in capital req. leads to ↓ in lending by 7% compared with the baseline.
Junge and Kugler (2012)	Swiss banks, 1999–2010. Calibration.	Halving leverage leads to ↑ in cost for nonfinancial sector by 0.6–1.5 bps and ↓ in permanent annual GDP by 0.04–0.05%.
Miles, Yang, and Marcheggiano (2013)	U.K. banks, 1997–2010. Calibration assuming half MM effect and one-third pass-through.	Halving leverage leads to ↑ in lending spread by 6 bps and ↓ in GDP by 0.15%.
Buch and Prieto (2014)	German banks, 1965–2009. Empirical estimation.	1 pp ↑ in capital req. leads to ↑ in bank loans by 0.22% in the long run.

(*continued*)

Table A.1. (Continued)

Paper	Data and Method	Cost of Capital
	A. Steady-State Impact	
Cohen and Scatigna (2014)	Large banks in advanced and emerging economies, 2009–12. Empirical estimation.	1 pp ↑ in capital ratio leads to ↑ in asset growth rate over the subsequent three years by 3 pp.
Martinez-Miera and Suarez (2014)	Dynamic general equilibrium model.	↑ in capital req. from 7% to 14% leads to ↑ in loan rate by 1.5 pp and ↓ in GDP by 8.5%, but also ↑ in aggregate net consumption of 0.9%.
Baker and Wurgler (2015)	U.S. banks, 1971–2011. Empirical estimation.	10 pp ↑ in tier 1 ratio leads to ↑ in cost of capital by 100–130 bps.
Brooke et al. (2015)	U.K. banks, 1997–2004. Calibration assuming MM effect and a full pass-through.	1 pp ↑ in tier 1 ratio leads to ↑ in lending rate by 5–10 bps and ↓ in permanent annual GDP by 0.01–0.05%.
Cline (2015)	U.S. banks, 2001–13. Calibration assuming half MM effect and a full pass-through.	15 pp ↑ in leverage ratio leads to ↑ in lending rate by 62 bps and ↓ in future path of GDP by 1.08 pp.
Kisin and Manela (2016)	U.S. banks, 2002–07. Empirical estimation.	1 pp ↑ in capital req. leads to ↑ in lending rate by 0.3 bps and ↓ in loan by 0.15%.
Begenau and Landvoigt (2017)	U.S. banks, 1999–2015. Dynamic general equilibrium model.	5 pp ↑ in capital req. leads to ↑ in lending by 3.6%, but has insignificant effect on GDP.
Clark, Jones, and Malmquist (2017)	U.S. banks, 1996–2012. Empirical estimation.	Halving leverage leads to ↑ in cost of capital by 14 bps for largest banks and ↓ in real GDP by 4.4 bps.
Dell'Ariccia, Laeven, and Suarez (2017)	U.S. banks, 1997–2011. Empirical estimation.	1 pp ↑ in capital req. leads to ↑ in lending by 0.6 pp.
Firestone, Lorenc, and Ranish (2017)	U.S. banks, 1988–2014. Calibration assuming half MM effect and a full pass-through.	1 pp ↑ in capital req. leads to ↑ in lending rate by 7 bps and ↓ in long-run GDP by 7.4 bps.

(*continued*)

Table A.1. (Continued)

Paper	Data and Method	Cost of Capital
	A. Steady-State Impact	
Barth and Miller (2018)	U.S. banks, 1892–2014. Empirical estimation.	↑ in leverage ratio from 4% to 15% leads to ↑ in cost of capital by 30 bps.
Gambacorta and Shin (2018)	International banks in advanced economies, 1994–2012. Empirical estimation.	1 pp ↑ in leverage ratio leads to ↓ in debt financing by 4 bps and ↑ in annual loan growth by 0.6 pp.
	B. Transitional Impact	
Furfine (2010)	U.S. banks, 1989–97. Dynamic general equilibrium model.	1 pp ↑ in capital req. leads to ↓ in lending growth rate by 5% in the first quarter.
Macroeconomic Assessment Group (2010)	Member countries using Financial Services Authority approach.	1 pp ↑ in capital req. implementation over eight years leads to, by 35th quarter, ↑ in lending spreads by 17 bps, ↓ in lending volume by 1.5%, and ↓ in GDP by 0.16% on average.
de Resende, Dib, and Perevalov (2010)	Canadian banks. Dynamic general equilibrium model.	6 pp ↑ in capital req. with a phase-in of four years leads to ↑ in lending spreads by 2 bps, ↓ in lending by almost 2%, ↓ in investment by 2.7%, and ↓ in GDP by 0.38%.
Institute of International Finance (2010)	G-3 banks, 2001–09. Calibration using a series of macro-banking-economic models.	2 pp ↑ in capital req. + other measures lead to, over the first five years, ↑ in lending spreads by 132 pp, ↓ in GDP growth by 0.6 pp per year on average.
Maurin and Toivanen (2012)	European banks, 2005–11. Empirical estimation.	1 pp capital gap dampens lending growth by 2–2.3 pp in median term.

(*continued*)

Table A.1. (Continued)

Paper	Data and Method	Cost of Capital
B. Transitional Impact		
de Ramon et al. (2012)	U.K. banks, 1992–2010. Empirical estimation.	1 pp ↑ in aggregate capital ratio leads to ↑ in lending spreads for household sector by 6.7 bps and for corporate sector by 19 bps.
Aiyar, Calomiris, and Wieladek (2014)	U.K. banks, 1998–2007. Empirical estimation.	1 pp ↑ in capital req. leads to cumulative ↓ in loan growth rate by 5.7 pp over the first three quarters.
Bridges et al. (2014)	U.K. banks, 1990–2011. Empirical estimation.	1 pp ↑ in capital req. leads to ↓ in household secured loans by 0.9 pp and commercial real estate lending by 8 pp over the first year. The first effect vanishes within three years, but the commercial real estate lending is reduced by 1.3 pp permanently.
Brun, Fraisse, and Thesmar (2014)	French banks, 2006–12. Empirical estimation.	1 pp ↓ in capital ratio leads to ↑ in loan size by 5 pp over the short period.
Corbae and D'Erasmo (2014)	U.S. banks, 2000–10. Dynamic general equilibrium model.	2 pp ↑ in capital req. leads to ↑ in loan interest rate by 50 bps and ↓ in loan supply by nearly 9%.
Mesonnier and Monks (2014)	European banks, 2011–12. Empirical estimation.	1 pp ↑ in capital req. leads to ↓ in annual loan growth rate by 1.2–1.6 pp over nine months.
Noss and Toffano (2016)	U.K. banks, 1986–2010. Empirical estimation.	15 pp ↑ in aggregate capital ratio leads to ↓ in lending by 1.4 pp and ↓ in GDP by 0.25% after 16 quarters.

References

Acharya, V., R. Engle, and M. Richardson. 2012. "Capital Shortfall: A New Approach to Ranking and Regulating Systemic Risks." *American Economic Review* 102 (3): 59–64.

Admati, A. R., P. M. DeMarzo, M. F. Hellwig, and P. C. Pfleiderer. 2010. "Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity is Not Expensive." Preprint of Max Planck Institute for Research on Collective Goods No. 2010/42.

Admati, A. R., and M. Hellwig. 2014. *The Bankers' New Clothes: What's Wrong with Banking and What to Do about It.* Princeton University Press.

Aiyar, S., C. Calomiris, and T. Wieladek. 2014. "Does Macroprudential Regulation Leak? Evidence from a UK Policy Experiment." *Journal of Money, Credit and Banking* 46 (s1): 181–214.

Allen, F., E. Carletti, and R. Marquez. 2011. "Credit Market Competition and Capital Regulation." *Review of Financial Studies* 24 (4): 983–1018.

———. 2015. "Deposits and Bank Capital Structure." *Journal of Financial Economics* 118 (3): 601–19.

Allen, F., and D. Gale. 2000. "Financial Contagion." *Journal of Political Economy* 108 (1): 1–33.

Avramova, S., and V. Le Lesle. 2012. "Revisiting Risk-Weighted Assets." IMF Working Paper No. 12/90.

Baker, M., and J. Wurgler. 2015. "Do Strict Capital Requirements Raise the Cost of Capital? Bank Regulation, Capital Structure, and the Low-Risk Anomaly." *American Economic Review: Papers and Proceedings* 105 (5): 315–20.

Bank for International Settlements. 2010. "An Assessment of the Long-Term Economic Impact of Stronger Capital and Liquidity Requirements." (August). http://www.bis.org/publ/bcbs173.pdf.

———. 2011. "Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems." Revised version (June). http://www.bis.org/publ/bcbs189.pdf.

———. 2013. "Global Systemically Important Banks: Updated Assessment Methodology and the Higher Loss Absorbency Requirement." (July). http://www.bis.org/publ/bcbs255.pdf.

———. 2016. "Standard TLAC Holdings." (July). http://www.bis.org/bcbs/publ/d387.pdf.

Bank of England. 2010. "Financial Stability Report." No. 27.

Barth, J. R., G. Caprio, and R. Levine. 2013. "Bank Regulation and Supervision in 180 Countries from 1999 to 2011." *Journal of Financial Economic Policy* 5 (2): 111–219.

Barth, J. R., and S. M. Miller. 2018. "Benefits and Costs of a Higher Bank 'Leverage Ratio'." *Journal of Financial Stability* 38 (October): 37–52.

Bebchuk, L. A., and I. Goldstein. 2011. "Self-fulfilling Credit Market Freezes." *Review of Financial Studies* 24 (11): 3519–55.

Begenau, J. 2014. "Capital Requirements, Risk Choice, and Liquidity Provision in a Business Cycle Model." Working Paper, Harvard Business School.

Begenau, J., and T. Landvoigt. 2017. "Financial Regulation in a Quantitative Model of the Modern Banking System." Working Paper, Harvard Business School.

Berrospide, J. 2013. "Bank Liquidity Hoarding and the Financial Crisis: An Empirical Evaluation." Finance and Economics Discussion Series No. 2013-03, Board of Governors of the Federal Reserve System.

Berrospide, J., and R. M. Edge. 2010. "The Effects of Bank Capital on Lending: What Do We Know, and What Does It Mean?" *International Journal of Central Banking* 6 (4, December): 5–54.

Board of Governors of the Federal Reserve System. 2015. "Dodd-Frank Act Stress Test 2015: Supervisory Stress Test Methodology and Results."

Borio, C. 2014. "Macroprudential Frameworks: (Too) Great Expectations?" *Central Banking Journal* (August).

Boyd, J., and G. De Nicolo. 2005. "The Theory of Bank Risk Taking and Competition Revisited." *Journal of Finance* 60 (3): 1329–43.

Bridges, J., D. Gregory, M. Nielsen, S. Pezzini, A. Radia, and M. Spaltro. 2014. "The Impact of Capital Requirements on Bank Lending." Working Paper No. 486, Bank of England.

Brooke, M., O. Bush, R. Edwards, J. Ellis, B. Francis, R. Harimohan, K. Neiss, and C. Siegert. 2015. "Measuring the Macroeconomic Costs and Benefits of Higher UK Bank Capital Requirements." Financial Stability Paper No. 35, Bank of England.

Brun, M., H. Fraisse, and D. Thesmar. 2014. "The Real Effects of Bank Capital Requirements." Mimeo, Banque de France.

Bryant, J. 1980. "A Model of Reserves, Bank Runs, and Deposit Insurance." *Journal of Banking and Finance* 4 (4): 335–44.

Buch, C. M., and E. Prieto. 2014. "Do Better Capitalized Banks Lend Less? Long-run Panel Evidence from Germany." *International Finance* 17 (1): 1–23.

Caballero, R. J., T. Hoshi, and A. K. Kashyap. 2008. "Zombie Lending and Depressed Restructuring in Japan." *American Economic Review* 98 (5): 1943–77.

———. 2009. "Global Imbalances and Financial Fragility." *American Economic Review* 99 (2): 584–88.

Caballero, R. J., and A. Krishnamurthy. 2008. "Collective Risk Management in a Flight to Quality Episode." *Journal of Finance* 63 (5): 2195–230.

Calem, P., and R. Rob. 1999. "The Impact of Capital-based Regulation on Bank Risk-taking." *Journal of Financial Intermediation* 8 (4): 317–52.

Calomiris, C. W., and C. M. Kahn. 1991. "The Role of Demandable Debt in Structuring Optimal Banking Arrangements." *American Economic Review* 81 (3): 497–513.

Cecchetti, S. 2014. "The Jury Is In." CEPR Policy Insight No. 76.

Claessens, S. 2015. "An Overview of Macroprudential Policy Tools." *Annual Review of Financial Economics* 7 (1): 397–422.

Claessens, S., S. R. Ghosh, and R. Mihet. 2013. "Macro-prudential Policies to Mitigate Financial System Vulnerabilities." *Journal of International Money and Finance* 39 (December): 153–85.

Claessens, S., and L. Ratnovski. 2014. "What Is Shadow Banking?" IMF Working Paper No. 14/25.

Clark, B., J. Jones, and D. Malmquist. 2017. "Leverage and the Cost of Capital for U.S. Banks." SSRN Working Paper.

Cline, W. R. 2015. "Testing the Modigliani-Miller Theorem of Capital Structure Irrelevance for Banks." Working Paper No. 15-8, Peterson Institute for International Economics.

———. 2016. "Benefits and Costs of Higher Capital Requirements for Banks." Working Paper No. 16-6, Peterson Institute for International Economics.

Cohen, B. H., and M. Scatigna. 2014. "Banks and Capital Requirements: Channels of Adjustment." BIS Working Paper No. 443.

Corbae, D., and P. D'Erasmo. 2014. "Capital Requirements in a Quantitative Model of Banking Industry Dynamics." Working Paper No. 14-13, Federal Reserve Bank of Philadelphia.

Cosimano, T. F., and D. S. Hakura. 2011. "Bank Behavior in Response to Basel III: A Cross-country Analysis." IMF Working Paper No. 11/119.

Dang, T. V., G. Gorton, B. Holmström, and G. Ordonez. 2017. "Banks as Secret Keepers." *American Economic Review* 107 (4): 1005–29.

DeAngelo, H., and R. M. Stulz. 2013. "Why High Leverage is Optimal for Banks." NBER Working Paper No. 19139.

De Mooji, R. 2011. "Tax Biases in Debt Finance: Assessing the Problem, Finding Solutions." IMF Staff Discussion Note No. 11/11.

de Ramon, S., Z. Iscenko, M. Osborne, M. Straughan, and P. Andrews. 2012. "Measuring the Impact of Prudential Policy on the Macroeconomy." FSA Occasional Paper No. 42.

de Resende, C., A. Dib, and N. Perevalov. 2010. "The Macroeconomic Implications of Changes in Bank Capital and Liquidity Requirements in Canada: Insights from the BoC-GEM-FIN." Discussion Paper No. 10-16, Bank of Canada.

Dell'Ariccia, G., L. Laeven, and G. Suarez. 2017. "Bank Leverage and Monetary Policy's Risk-taking Channel; Evidence from the United States." *Journal of Finance* 72 (2): 613–54.

Dell'Ariccia, G., and L. Ratnovski. 2013. "Bailouts and Systemic Insurance." IMF Working Paper No. 13/233.

Diamond, D. W., and P. H. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (3): 401–19.

Diamond, D. W., and R. G. Rajan. 2000. "A Theory of Bank Capital." *Journal of Finance* 55 (6): 2431–65.

Eber, M., and C. Minoiu. 2015. "Tightening the Screws: How Do Banks Adjust to Stricter Supervision?" IMF Working Paper, forthcoming.

Elliott, D. J. 2009. "Quantifying the Lending Effects of Increased Capital Requirements." The Brookings Institution.

Enrich, D., and M. Colchester. 2012. "EU Banks' Risk in Eyes of Beholder." *Wall Street Journal*, June 22.

Esty, B. 1998. "The Impact of Contingent Liability on Commercial Bank Risk Taking." *Journal of Financial Economics* 47 (2): 189–218.

European Banking Authority. 2014. "Results of 2014 EU-Wide Stress Test." Report (October 26).

Federal Reserve Bank of Minneapolis. 2017. "The Minneapolis Plan to End Too Big to Fail." (December).

Firestone, S., A. Lorenc, and B. Ranish. 2017. "An Empirical Economic Assessment of the Costs and Benefits of Bank Capital in the US." Finance and Economics Discussion Series No. 2017-034, Board of Governors of the Federal Reserve System.

Flannery, M. 2014. "Contingent Capital Instruments for Large Financial Institutions: A Review of the Literature." *Annual Review of Financial Economics* 6: 225–40.

Francis, W. B., and M. Osborne. 2012. "Capital Requirements and Bank Behavior in the UK: Are There Lessons for International Capital Standards?" *Journal of Banking and Finance* 36 (3): 803–16.

Fratianni, M., and Marchionne. 2013. "The Fading Stock Market Response to Announcements of Bank Bailouts." *Journal of Financial Stability* 9 (1): 69–89.

Furfine, C. 2010. "Evidence on the Response of US Banks to Changes in Capital Requirement." BIS Working Paper No. 88.

Gale, D., and O. Özgür. 2005. "Are Bank Capital Ratios Too High or Too Low? Incomplete Markets and Optimal Capital Structure." *Journal of the European Economic Association* 3 (2–3): 690–700.

Gambacorta, L., and H. S. Shin. 2018. "Why Bank Capital Matters for Monetary Policy." *Journal of Financial Intermediation* 35 (Part B): 17–29.

Geithner, T. 2014. *Stress Test: Reflections on Financial Crises.* New York: Random House.

Gennaioli, N., A. Shleifer, and R. Vishny. 2012. "Neglected Risks, Financial Innovation, and Financial Fragility." *Journal of Financial Economics* 104 (3): 452–68.

Gennaioli, N., A. Shleifer, and R. W. Vishny. 2013. "A Model of Shadow Banking." *Journal of Finance* 68 (4): 1331–63.

Goodhart, C. 2010. "How Should We Regulate Bank Capital and Financial Products? What Role for 'Living Wills'?" In *The*

Future of Finance: The LSE Report, 165–86 (chapter 5). London School of Economics and Political Science.

Gornall, W., and I. A. Strebulayev. 2018. "Financing as a Supply Chain: The Capital Structure of Banks and Borrowers." *Journal of Financial Economics* 129 (3): 510–30.

Gorton, G., S. Lewellen, and A. Metrick. 2012. "The Safe-Asset Share." *American Economic Review* 102 (3): 101–6.

Gorton, G., and A. Metrick. 2012. "Securitized Banking and the Run on Repo." *Journal of Financial Economics* 104 (3): 425–51.

Gorton, G. B., and G. Ordonez. 2013. "The Supply and Demand for Safe Assets." NBER Working Paper No. 18732.

Gorton, G., and G. Pennacchi. 1990. "Financial Intermediaries and Liquidity Creation." *Journal of Finance* 45 (1): 49–71.

Greenwood, R. M., S. G. Hanson, and J. C. Stein. 2015. "A Comparative-Advantage Approach to Government Debt Maturity." *Journal of Finance* 70 (4): 1683–1722.

Hakenes, H., and I. Schnabel. 2011. "Capital Regulation, Bank Competition, and Financial Stability." *Economics Letters* 113 (3): 256–58.

Hellmann, T. F., K. C. Murdock, and J. E. Stiglitz. 2000. "Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?" *American Economic Review* 90 (1): 147–65.

Hellwig, M. F. 2014. "Liquidity Provision and Equity Funding of Banks." Working Paper.

Holmstrom, B., and J. Tirole. 1997. "Financial Intermediation, Loanable Funds, and the Real Sector." *Quarterly Journal of Economics* 112 (3): 663–91.

Huang, R., and L. Ratnovski. 2011. "The Dark Side of Bank Wholesale Funding." *Journal of Financial Intermediation* 20 (2): 248–63.

Institute of International Finance. 2010. "Interim Report on the Cumulative Impact on the Global Economy of Proposed Changes in the Banking Regulatory Framework." (June). http://ebf-fbe.eu/uploads/10-Interim%20NCI_June2010_Web.pdf.

Johnston Ross, E. B., and L. Shibut. 2015. "What Drives Loss Given Default? Evidence from Commercial Real Estate Loans at Failed Banks." Working Paper No. 2015-03, FDIC Center for Financial Research.

Junge, G., and P. Kugler. 2012. "Quantifying the Impact of Higher Capital Requirements on the Swiss Economy." WWZ Discussion Paper No. 2012/13.

Kane, E. J. 1989. *The S & L Insurance Mess: How Did It Happen?* Urban Institute Press.

Kashyap, A. K., R. G. Rajan, and J. C. Stein. 2008. "Rethinking Capital Regulation." In *Maintaining Stability in a Changing Financial System*. Proceedings of the 2008 Economic Policy Symposium sponsored by the Federal Reserve Bank of Kansas City. Kansas, City, MO: Federal Reserve Bank of Kansas City.

Kashyap, A. K., J. C. Stein, and S. Hanson. 2010. "An Analysis of the Impact of 'Substantially Heightened' Capital Requirements on Large Financial Institutions." Working Paper, Harvard University.

Keeley, M. C. 1990. "Deposit Insurance, Risk, and Market Power in Banking." *American Economic Review* 80 (5): 1183–1200.

King, M. 2010. "Mapping Capital and Liquidity Requirements to Bank Lending Spreads." BIS Working Paper No. 324 (November).

Kisin, R., and A. Manela. 2016. "The Shadow Cost of Bank Capital Requirements." *Review of Financial Studies* 29 (7): 1780–1820.

Krishnamurthy, A. 2010. "How Debt Markets Have Malfunctioned in the Crisis." *Journal of Economic Perspectives* 24 (1): 3–28.

Krishnamurthy, A., and A. Vissing-Jorgensen. 2012. "The Aggregate Demand for Treasury Debt." *Journal of Political Economy* 120 (2): 233–67.

Laeven, L., and R. Levine. 2009. "Bank Governance, Regulation and Risk Taking." *Journal of Financial Economics* 93 (2): 259–75.

Laeven, L., and L. Ratnovski. 2014. "Corporate Governance of Banks and Financial Stability." VoxEU column. http://www.voxeu.org/article/corporate-governance-banks-and-financial-stability.

Laeven, L., and F. Valencia. 2018. "Systemic Banking Crises Revisited." IMF Working Paper No. 18/206.

Macroeconomic Assessment Group. 2010. "Assessing the Macroeconomic Impact of the Transition to Stronger Capital and Liquidity Requirements." Bank for International Settlements. http://www.bis.org/publ/othp12.pdf.

Marcus, A. J. 1984. "Deregulation and Bank Financial Policy." *Journal of Banking and Finance* 8 (4): 557–65.

Mariathasan, M., and O. Merrouche. 2014. "The Manipulation of Basel Risk-weights." *Journal of Financial Intermediation* 23 (3): 300–21.

Martin, A., and B. M. Parigi. 2013. "Bank Capital Regulation and Structured Finance." *Journal of Money, Credit and Banking* 45 (1): 87–119.

Martinez-Miera, D., and J. Suarez. 2014. "Banks' Endogenous Systemic Risk Taking." Mimeo, CEMFI.

Matutes, C., and X. Vives. 2000. "Imperfect Competition, Risk Taking, and Regulation in Banking." *European Economic Review* 44 (1): 1–34.

Maurin, L., and M. Toivanen. 2012. "Risk, Capital Buffer and Bank Lending: A Granular Approach to the Adjustment of Euro Area Banks." ECB Working Paper No. 1499.

Mendicino, C., K. Nikolov, J. Suarez, and D. Supera. 2015. "Welfare Analysis of Implementable Macroprudential Policy Rules: Heterogeneity and Trade-offs." Mimeo, European Central Bank and CEMFI.

Mesonnier, J., and A. Monks. 2014. "Did the EBA Capital Exercise Cause a Credit Crunch in the Euro Area?" Working Paper, Bank of France.

Miles, D., J. Yang, and G. Marcheggiano. 2013. "Optimal Bank Capital." *Economic Journal* 123 (567, March): 1–37.

Modigliani, F., and M. H. Miller. 1958. "The Cost of Capital, Corporation Finance and the Theory of Investment." *American Economic Review* 48 (3): 261–97.

Myers, S. C., and N. S. Majluf. 1984. "Corporate Financing and Investment Decisions when Firms Have Information that Investors Do Not Have." *Journal of Financial Economics* 13 (2): 187–221.

Nguyen, T. T. 2013. "Bank Capital Requirements: A Quantitative Analysis." Working Paper, Ohio State University.

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.

Peek, J., and E. Rosengren. 2000. "Collateral Damage: Effects of the Japanese Bank Crisis on Real Activity in the United States." *American Economic Review* 90 (1): 30–45.

Perotti, E., L. Ratnovski, and R. Vlahu. 2011. "Capital Regulation and Tail Risk." *International Journal of Central Banking* 7 (4, December): 123–63.

Plantin, G. 2015. "Shadow Banking and Bank Capital Regulation." *Review of Financial Studies* 28 (1): 146–75.

Ratnovski, L. 2013. "How Much Capital Should Banks Have?" VOX column. http://www.voxeu.org.

Repullo, R. 2004. "Capital Requirements, Market Power, and Risk-taking in Banking." *Journal of Financial Intermediation* 13 (2): 156–82.

Santos, J. A. C., and A. Winton. 2010. "Bank Capital, Borrower Power, and Loan Rates." Paper presented at the 2010 American Finance Association Annual Meeting, Atlanta, Georgia, USA, January 3–5.

Saurina, J. 2009. "Loan Loss Provisions in Spain: A Working Macroprudential Tool." *Revista de Estabilidad Financiera* (Banco de España) 17 (November): 11–26.

Schanz, J., D. Aikman, P. Collazos, M. Farag, D. Gregory, and S. Kapadia. 2011. "The Long-term Economic Impact of Higher Capital Levels." *BIS Papers* 60: 73–81.

Schuermann, T. 2004. "What Do We Know about Loss Given Default?" Mimeo, Federal Reserve Bank of New York.

Shibut, L., and R. Singer. 2014. "Loss Given Default for Commercial Loans at Failed Banks." Mimeo, Federal Deposit Insurance Corporation.

Slovik, P., and P. Cournède. 2011. "Macroeconomic Impact of Basel III." OECD Economics Department Working Paper No. 844.

Song, F., and A. V. Thakor. 2007. "Relationship Banking, Fragility, and the Asset-Liability Matching Problem." *Review of Financial Studies* 20 (6): 2129–77.

Swagel, P. 2009. "The Financial Crisis: An Inside View." *Brookings Papers on Economic Activity* 40 (1, Spring): 1–78.

Van den Heuvel, S. J. 2008. "The Welfare Cost of Bank Capital Requirements." *Journal of Monetary Economics* 55 (2): 298–320.

Zhou, J., V. Rutledge, W. Bossu, M. Dobler, N. Jassaud, and M. Moore. 2012. "From Bail-out to Bail-in: Mandatory Debt Restructuring of Systemic Financial Institutions." IMF Staff Discussion Note No. 12/03 (April 24).

Relaxed Lending Standards and the 2007 Mortgage Crisis: Changes in Household Debt and Borrowing Behaviors*

Seda Durguner
University of Southern California

Using the 1998 and 2007 Survey of Consumer Finances data, this paper examines changes in household debt and borrowing behaviors following a decline in lending standards. The findings suggest that households desired more debt and obtained more of their desired debt in 2007 than they did in 1998. This is especially true for the credit-constrained households headed by someone that was 34 years of age or younger or between 54 and 65, and for constrained households with annual income of between $30,000 and $60,000 or of more than $345,000. Households headed by someone that was 34 or younger or between 44 and 65, and households with income levels of more than $60,000 were also less likely to be credit constrained in 2007.

JEL Codes: G01, G21, G28, D12, D14, D91.

1. Introduction

Research in the wake of the 2007 subprime mortgage crisis has found that its contributing factors included a decline in lending standards and a rise in household borrowing, especially by higher-risk borrowers who drew against their rising home equity.[1] Although the decline in lending standards has been well documented in literature,

*Author contact: 701 Exposition Blvd., Ste. 231. HOH-231, MC-1422. Los Angeles, California 90089-1422, USA. E-mail: sdurgune@usc.edu. Tel: 213 740 6515.

[1]For the decline in lending standards, see Dell'Ariccia, Igan, and Laeven (2008), Demyanyk and Van Hemert (2011), Foote et al. (2008), Keys et al. (2009, 2010), Mian and Sufi (2009b), and Purnanandam (2011). For the rise in household borrowing, see Mian and Sufi (2009a, 2010).

the change in household debt and borrowing behavior has not been analyzed in light of the decline in lending standards. This paper adds to the literature on the financial crisis by examining in detail the changes in household desired debt and whether households were able to obtain a higher proportion of their desired debt in the run-up to 2007. The goal is not to show the generalized trends of more borrowing since 1989. The goal is to show whether households had a greater capacity to borrow following the decline in lending standards. Through use of the Federal Reserve Board's 1998 and 2007 Survey of Consumer Finances (SCF), this paper addresses three questions regarding the relaxation of lending standards in the 1998–2007 period: (i) Did household demand for credit rise? (ii) Were households less likely to be credit constrained (that is, more likely to get all the credit they desired)? (iii) Did households that were credit constrained obtain more of the credit they desired? Households are subgrouped into different categories based on their age and income. The paper analyzes how answers to the three questions vary among households in different age and income subgroups.

The literature on the role of lending standards in the 2007 subprime mortgage crisis has focused on mortgage loan characteristics and mortgage loan qualities to draw inferences about lending standards. The goal in this paper is to analyze more directly how households' desired debt and borrowing behaviors changed in a period of declining lending standards. The focus is on households generally, not just mortgage borrowers, because the unemployment and loss of wealth caused by the crisis were similarly widespread. Conducted every three years, the SCF is a rigorous, representative sampling of U.S. households that provides detailed information about their outstanding debt levels, financial and nonfinancial assets, attitudes about debt, and expectations about the economy.[2] This study is the first to measure households' desired debt in assessing their borrowing behavior in the years leading up to the 2007 crisis.[3]

[2]See http://www.federalreserve.gov/econresdata/scf/scfindex.htm.

[3]Desired debt has been studied in other contexts. See, for example, Cox and Jappelli (1993), Crook (2001), Crook and Hochguertel (2006), Duca and Rosenthal (1993), Gropp, Scholz, and White (1997), Lyons (2003), Magri (2007), and Manrique and Ojah (2004).

As part of its analysis, this study computes, for various household characteristics in both 1998 and 2007, the amount of debt desired by credit-constrained households and the gap between their desired and actual stock of debt; it then examines how this gap changed from 1998 to 2007 for various characteristics of credit-constrained households. In 1998, lenders were making subprime mortgages, but the decline in lending standards had not yet begun; by 2007, lending standards had been relaxed for some time. Any difference in household debt behavior between 1998 and 2007 will be the more notable because household expectations about interest rates and the general economy, as documented in the surveys, were similar across the two years.[4]

Following the decline in lending standards, one would expect that credit-constrained households will obtain more of their desired debt and that the gap between actual and desired stock of debt will decline. Thus, as a means of analyzing the changes in households' debt and borrowing behaviors, this study tests the following null hypothesis: The gap between the desired and actual stock of debt for credit-constrained households was lower in 2007 than in 1998. Testing for this hypothesis will help address the three issues posed at the outset: whether the amount of debt desired rose between the two years, whether households became less likely to be credit constrained, and whether credit-constrained households in 2007 received more of their desired debt than in 1998. It will also be possible to address which households—and specifically which age and income groups—have experienced those changes.

To identify credit-constrained households, the present study uses the self-reported indicators in the SCF data that have been widely used in the literature (Cox and Jappelli 1990, 1993; Crook 1996, 2001; Jappelli 1990; Lyons 2003): whether the household has ever

[4]The interviews for the 1998 SCF were conducted between July and December 1998; those for the 2007 SCF were conducted between May and December 2007. Regarding attitudes about the economy in the 1998 (2007) survey, 48 percent (41 percent) of households expected the economy to be the same in the next five years, 24 percent (29 percent) expected it to be better, and 28 percent (30 percent) expected it to be worse. Regarding attitudes about interest rates in the 1998 (2007) survey, 63 percent (63 percent) of households expected them to be lower five years from now, 30 percent (29 percent) expected them to be the same, and 7 percent (8 percent) expected them to be higher.

been denied credit, or has ever been given less credit than it applied for, or has ever been discouraged from applying for credit. Because these self-reported constraint indicators are for any loan type, the analysis here encompasses all household debt.

One difficulty with testing the null hypothesis is determining the desired stock of debt of credit-constrained households. The stock of debt desired by unconstrained households can be taken to be the same as their outstanding debt, as reported in the SCF, because they can obtain the full loan amount they desire. However, households with borrowing constraints can obtain only part of the loan amount they desire. The SCF data show neither the loan amounts originally requested by constrained households nor the amounts that they did not apply for because they assumed that the request would be denied. Thus, the analysis must estimate the stock of debt desired by constrained households.

To do so, this study follows the procedure in Cox and Jappelli (1993). As an initial step, the desired stock of debt for all unconstrained households is estimated and derived from indebted unconstrained households. As the second step, the desired stock of debt for the constrained households is derived from the coefficients estimated in the initial step.

The literature on household debt has not considered the entire distribution of net wealth that is provided in the SCF data. This study does so; it includes wealthy households in its analysis and adjusts for the SCF's oversampling of wealthy households by using the sampling weights provided in the SCF data. Other ways in which this study extends the literature is that it applies previously used techniques—constructs of permanent income and instrumenting the net worth variable used in the regressions—to more recent data (the 2007 SCF).[5] The standard errors are also bootstrapped to generate correct standard errors.

The findings of the present analysis show that desired debt was greater in 2007 than in 1998. And even though desired debt rose, credit-constrained households were able to receive more of their

[5]The construction of permanent income follows the techniques developed in Cox and Jappelli (1993) and King and Dicks-Mireaux (1981). The regressions used here control for constructed permanent income, and the standard errors are bootstrapped to generate correct standard errors.

desired debt. This was especially true for households with household heads that were age 34 or younger or between 54 and 65, with annual income of between $30,000 and $60,000, and with income of more than $345,000. And less likely to be credit constrained in 2007 were households with income of more than $60,000 and with household heads that were age 34 or younger or between 44 and 65.

These findings imply that households had a greater capacity to borrow following the decline in lending standards. Households desired more debt and were less likely to be credit constrained, and credit-constrained households received more of their desired debt. These results warrant the greater attention that regulators are placing on credit markets to preserve financial stability. Monitoring the status of lending standards, desired debt, and credit constraints in time allows regulators to apply precautionary restraints across credit markets to better protect borrowers, the financial system, and the wider economy.

The paper proceeds as follows. Sections 2, 3, and 4 review the literature on household debt and present the theory and model. Sections 5, 6, and 7 describe the variables, cover the data and descriptive statistics, and present the findings. Section 8 concludes.

2. Literature Review

Many papers have attempted to determine whether households are constrained in their access to credit, and most of them used SCF data for their analysis. To determine the presence of constraint, some used indirect measures such as assets, the wealth-to-income ratio, and the saving rate. The findings of the papers using indirect measures have been ambiguous. Some of them concluded that approximately 20 percent of U.S. households had credit constraints (Hall and Mishkin 1982; Hayashi 1985; Hubbard and Judd 1986; Mariger 1987; Zeldes 1989). Other studies (Altonji and Siow 1987; Attanasio and Weber 1995; Dejuan and Seater 1999; Runkle 1991) with indirect measures concluded that households did not have any credit constraints. Lyons (2003) suggested that the ambiguity of these findings was probably due to their lack of a more direct measure to identify households with credit constraints.

A large number of other studies (Cox and Jappelli 1990, 1993; Crook 1996, 2001; Duca and Rosenthal 1993; Fissell and Jappelli

1990; Gropp, Scholz, and White 1997; Jappelli 1990; Lyons 2003) have used a more direct measure from the SCF data, namely, self-reporting on the questions regarding constrained access: whether the household (i) has ever been denied credit, (ii) has ever been given less credit than the amount it applied for, and (iii) has ever been discouraged from applying for credit. Overall, Crook (1996) has found that 19 percent to 20 percent of households were credit constrained prior to 1983 and 1989. The studies found a wide range of factors that significantly influenced the probability of a household being credit constrained. Among the factors lowering the probability were a smaller household size and a household head with any of the following characteristics: older; white; married; a good credit history; a checking account; higher income; higher net worth; more numbers of debit, credit, and charge cards; owning a home; a longer tenure as a homeowner at the current address; a longer tenure as an employee at the current employer; more years of schooling; in the habit of saving; not receiving any welfare payments; and no near-term prospect of a major expenditure. Also decreasing the likelihood of a credit constraint was living in a more rural area and in an area with a lower level of banking competition. Findings by Gropp, Scholz, and White (1997) suggest that bankruptcy exemption, level of Herfindahl index for financial institutions in the area, and whether there are multibank holding companies in the state are also significant determinants of having credit constraints.

In a departure from the rest of the credit-constraint literature, Cox and Jappelli (1993) and Crook and Hochguertel (2006) related the determinants of credit constraint to an estimate of households' permanent income (that is, the long-term income households anticipate having on the basis of their assets, skills, and education), computed as in King and Dicks-Mireaux (1982). Cox and Jappelli (1993) find that being credit constrained is negatively related to permanent income at the weakest (10 percent) significance level. Among the Crook and Hochguertel (2006) findings for the United States is that the greater the amount by which current income exceeds permanent income, the higher is the probability of the household having a credit constraint. Findings by Crook and Hochguertel (2006) also show that the probability of constraint varies directly with household head's being self-employed, years of education, and the number of children. The probability varies inversely with net worth, income, household head's age, and household head's being unemployed, female, or single.

Some studies derived determinants of credit constraint as a step in estimating desired debt (Cox and Jappelli 1993; Crook 2001; Crook and Hochguertel 2006; Duca and Rosenthal 1993; Gropp, Scholz, and White 1997; Magri 2007; Manrique and Ojah 2004). They conditioned their estimates of desired debt on households that have positive amounts of desired debt and are not credit constrained. These studies found a wide range of factors that significantly influenced the desired debt. Among the factors increasing the desired debt were larger household size and a household head with any of the following characteristics: male, young, more years of schooling,[6] currently working, ownership of a home, higher net worth,[7] higher income,[8] willingness to borrow to finance luxury items, and near-term prospect of a major expenditure.

Cox and Jappelli (1993) showed that desired debt has a positive relation with permanent income and that desired debt declines after a threshold age. Crook and Hochguertel (2006) suggest that, in the United States, a reduction in current income below permanent income positively affects the amount of debt that is demanded. Rising age (until 30) also has a positive effect on desired debt, but desired debt levels decline after age 65. Crook (2001) showed a negative relation with age for household heads aged above 54 years old, income squared, and risk aversion. Gropp, Scholz, and White (1997) found a negative relation of desired debt to concentration of the financial market, living in certain regions, number of years the household head has been working at current employer, and to rising age for those between 34 and 65.

The literature excludes wealthy households from analysis and does not consider the entire distribution of net wealth that is provided with the SCF data. This study fills this gap by including wealthy households in the analysis and using SCF sampling weights. Use of the weights adjusts for the oversampling of wealthy

[6] Crook and Hochguertel (2006) find a negative relation with desired debt and education.

[7] Crook (2001) showed that desired debt declines with higher net worth.

[8] Cox and Jappelli (1993) found a negative relation with current income.

households and reflects the U.S. population. As another extension to the literature, this study includes more recent data (the 2007 SCF).

Simultaneity exists between net worth and desired debt, and between net worth and being credit constrained. It is not obvious whether the literature corrects for this simultaneity—see Crook (2001), Crook and Hochguertel (2006), and Manrique and Ojah (2004). This paper corrects for the simultaneity by instrumenting net worth in the manner of Cox and Jappelli (1993) and Duca and Rosenthal (1993).[9]

Crook (2005) suggests using permanent income while determining debt demand functions; few have done so and controlled for it—see Cox and Jappelli (1993), Lyons (2003), and Crook and Hochguertel (2006). This study constructs permanent income, and the regressions control for it; standard errors are bootstrapped to generate correct standard errors.

3. Theory

According to the life-cycle permanent-income hypothesis (LCPIH), households smooth their consumption over their lifetime in line with their expected lifetime (permanent) income. Households that anticipate an increase in their expected future incomes will borrow to increase their current consumption without waiting for the income increase to take effect. Without the ability to borrow, the LCPIH fails: households' current consumption levels are limited to their current income levels because they must defer any increase in consumption until their present income actually increases.

Following Zeldes (1989), the present study applies the LCPIH model to consumers operating with a borrowing constraint over two periods. Each consumer starts period 1 with some financial wealth (A_1), receives labor income at the beginning of each period, and chooses a consumption level. By the end of period 2, the consumers have spent all their financial and labor income and die without

[9]In estimating the desired amount of debt, Magri (2007) took into consideration the fact that net wealth can be affected by simultaneity problems; as a robustness check, she therefore instrumented net wealth with a lag of net wealth. However, she did not correct for this simultaneity in estimating the demand for debt or the probability of being unconstrained.

any debts or bequests. In this setting, the consumers solve for the problem given below to determine their optimal consumption levels that will maximize their total utility. The desired consumption level determines whether debt is desired in period 1 and how much is desired (D_1^*).

$$Max_{C_1 C_2} U(C_1) + U(C_2) \tag{1}$$

subject to

1. $A_2 = (A_1 + Y_1 - C_1)(1 + r)$
2. $A_2 + Y_2 = C_2$
3. $A_1 + Y_1 - C_1 \geq -BC$
4. $C_1, C_2 \geq 0.$

Utility (U), a one-period utility function, is assumed to be a constant relative risk aversion utility function defined as $U(C) = \frac{C^{1-\alpha}-1}{1-\alpha}$. C_1 and C_2 and Y_1 and Y_2 are consumption and labor income in each of the two periods; A_1 and A_2 are financial wealth at the beginning of each of the two periods (before labor income and consumption); r is the interest rate earned on holding funds from period 1 into period 2; and BC, the borrowing constraint, is the maximum amount that the consumer can borrow in period 1. The above problem can be reduced to the following maximization problem:[10]

$$Max_{C_1} F(C_1) = U(C_1) + U[(1 + r)(A_1 + Y_1 - C_1) + Y_2] \tag{2}$$

subject to

5. $C_1 \geq A_1 + Y_1 + \min\left(BC, \frac{Y_2}{1+r}\right)$
6. $C_1 \leq 0.$

[10] In equation (1), the combination of constraints 4, 2, and 1 gives $\mathrm{C_2} = (1 + \mathrm{r})(\mathrm{A_1} + \mathrm{Y_1} - \mathrm{C_1}) + \mathrm{Y_2} \geq 0$. Solving for $\mathrm{C_1}$ gives $\mathrm{C_1} \leq \mathrm{A_1} + \mathrm{Y_1} + \frac{\mathrm{Y_2}}{1+\mathrm{r}}$. Constraint 3 is the same as $C_1 \leq A_1 + Y_1 + BC$. The combination of the two inequalities for $\mathrm{C_1}$ gives $C_1 \leq \mathrm{A_1} + Y_1 + \min\left(BC, \frac{Y_2}{1+r}\right)$.

If there is no boundary solution, the solution to the maximization problem is

$$C_1^* = \frac{(1+r)^{-1/\alpha}(Y_2 + (1+r)(A_1 + Y_1))}{1 + (1+r)^{1-\frac{1}{\alpha}}}.$$

If there is a boundary solution, $C_1^* = A_1 + Y_1 + \min\left(BC, \frac{Y_2}{1+r}\right)$.

Appendix A shows the solution to the maximization problem. Given the desired consumption level, C_1^*, and the resources that are available (A_1, Y_1), the consumer determines the desired loan size and thus the desired stock of debt, D_1^*.

4. Model

To assess changes in household debt and borrowing behaviors between 1998 and 2007, this study tests the following null hypothesis: The gap between the desired and actual stock of household debt was lower in 2007 than in 1998. The test will help determine (i) whether households in 2007 desired more debt than households in 1998, (ii) whether households were less likely to be credit constrained in 2007 than in 1998, and (iii) whether credit-constrained households obtained a higher proportion of their desired debt in 2007 than in 1998.

The actual stock of debt is the debt outstanding at the time of the survey and is observed in the SCF data. Desired debt is not observed. The desired stock of debt for unconstrained households (D_{1U}^*) can be considered to have the same value as their actual stock of debt (D_{1U}) because unconstrained households can obtain the full amount of the loans they desire. For households with credit constraints, the actual stock of debt (D_{1C}) is by definition less than the desired stock (D_{1C}^*). Therefore, the debt desired by credit-constrained households must be estimated.

The gap between desired and actual debt for credit-constrained households is constructed in three steps for 1998 and again for 2007, following Cox and Jappelli (1993). In the first step for each year, the desired debt of households that have no credit constraints and have positive demand for debt is estimated on the basis of outstanding debt. In the second step, the desired stock of debt for credit-constrained households is computed with the estimated coefficients obtained from the first step. In the third step, the gap for credit-constrained households is calculated by measuring the difference between the estimated level of desired debt and the outstanding debt. The gap levels in 2007 are then compared with the gap levels in 1998.

The first step—estimating the desired stock of debt for the indebted and unconstrained households—consists of the generalized Tobit model shown in equations (3), (4), and (5). The main goal is to predict equation (3), the desired stock of debt for households having positive demand for debt and no credit constraint.

$$D_1^* = \beta_1 X_1 + \theta_1 \lambda_1 + \theta_2 \lambda_2 + \varepsilon_1, \tag{3}$$

where the dependent variable, D_1^*, is the desired stock of debt; D_1^* is observable only when D_1^* equals the amount of debt that is outstanding. This equality is true only when the household has a positive demand for debt and does not have any credit constraints. X_1 is a vector of variables that determine the desired consumption levels and the resources that are available to pay for consumption; hence, the vector implies the level of desired debt. The variables for the vector are net worth, family income, permanent income (constructed), homeownership, future expenses, attitudes toward debt, and demographics (as in Cox and Jappelli 1993; Crook 1996, 2001; Duca and Rosenthal 1993; Jappelli 1990).

Equations (4) and (5) are the selection equations needed for equation (3). Equation (4) defines the probability of having positive demand for debt, and equation (5) defines the probability of not having any credit constraints.

$$l_1 = \beta_2 X_2 + \varepsilon_2, \tag{4}$$

where l_1 is an unobserved or latent variable representing demand for debt. The observed variable, H, is defined by the following two relations:

- If $l_1 > 0$, then H = 1 and the household has outstanding debt.
- If $l_1 \leq 0$, then H = 0 and the household has no outstanding debt.

X_2 is a vector of variables that determine whether a household has a demand for debt. As in Crook (2001) and Duca and Rosenthal (1993), the same variables determine the existence of a demand for debt and the level of desired debt. Thus, the vector of X_2 is the same as the vector of X_1.

$$l_2 = \beta_3 X_3 + \varepsilon_3, \tag{5}$$

where l_2 is an unobserved or latent variable representing excess supply of credit. The observed variable, V, is defined by the following two relations:

- If $l_2 > 0$, then V = 1 and the household has no credit constraint.
- If $l_2 \leq 0$, then V = 0 and the household has a credit constraint.

Households will have credit constraints only when they demand more debt than lenders will grant them, which represents an excess demand for debt. Thus, X_3 is a vector of (i) variables that determine the existence of a demand for debt and the level of desired debt (the demand side) and (ii) of borrower characteristics that are used in credit-scoring models (the supply side) to guide lenders' decisions on how much credit to grant an applicant (as in Cox and Jappelli 1993; Crook 1996, 2001; Duca and Rosenthal 1993; Jappelli 1990). X_3 consists of X_2 and additional variables that proxy for credit history.[11]

The error terms are assumed to have a trivariate normal distribution, and ρ_{23} is assumed to be zero (as in Cox and Jappelli 1993; Crook 2001; Crook and Hochguertel 2006; Duca and Rosenthal 1993; Gropp, Scholz, and White 1997; Lyons 2003).

$$\varepsilon_1 \sim \mathrm{N}(0, \sigma^2); \varepsilon_2 \sim \mathrm{N}(0, 1); \varepsilon_3 \sim \mathrm{N}(0, 1)$$
$$\mathrm{corr}(\varepsilon_1, \varepsilon_2) = \rho_{12}; \mathrm{corr}(\varepsilon_1, \varepsilon_3) = \rho_{13}; \mathrm{corr}(\varepsilon_2, \varepsilon_3) = \rho_{23} = 0$$

[11] Cox and Jappelli (1993) and Gropp, Scholz, and White (1997) use regional dummies to control for variations in the regulation and characteristics of credit markets; this study does not do so because this information is not in the publicly available SCF data.

As in Duca and Rosenthal (1993), when $\rho_{23} = 0$, the estimate of equation (3) (the desired stock of debt) is derived from the estimate of equations (4) and (5) as follows:

- Estimate β_2 and β_3 by running a univariate probit on equations (4) and (5).
- Compute the inverse Mills ratios λ_1 and λ_2 and evaluate them at $\beta_2 X_2$ and $\beta_3 X_3$, respectively.[12]
- Include λ_1 and λ_2 in equation (3) as additional explanatory variables, and conduct an ordinary least squares (OLS) calculation on this regression for a sample of unconstrained households that have outstanding debt.

As shown in Duca and Rosenthal (1993), standard errors for this procedure will not be correct even though the procedure produces consistent estimates of β_1. Correct standard errors are produced through bootstrapping, which is done here with the replicate weights provided in the SCF data.

A similar issue arises with permanent income, an independent variable constructed here according to Cox and Jappelli (1993) and King and Dicks-Mireaux (1981). The estimation process for permanent income generates a measurement error, and such errors for an independent variable can raise endogeneity and inconsistent OLS estimates. Because the literature on household debt did not raise any endogeneity issues for the constructed permanent-income variable, this study assumes that it is exogenous—see Cox and Jappelli (1993), Crook and Hochguertel (2006), and Lyons (2003). According to Wooldridge (2002, p. 74), under exogeneity, the OLS procedure will produce consistent estimates but will inflate the error variance and thus inflate the variance for the parameter estimates and underestimate significance of the parameter estimates. Here again, bootstrapping is applied to generate correct standard errors.

Another issue is the simultaneity between net worth and debt. As noted in Duca and Rosenthal (1993), borrowing to finance the consumption of nondurables immediately lowers net worth; hence,

[12] The inverse Mills ratio for selection equation (4), the probability of demand for debt, is $\frac{f(\beta_2 X_2)}{F(\beta_2 X_2)}$, where $f(\ )$ is the probability density function and $F(\ )$ is the cumulative distribution function.

there is a simultaneous relationship between net worth and demand for debt and between net worth and desired debt level. Moreover, the observed level of net worth is sensitive to whether the household is credit constrained. Therefore, as in Cox and Jappelli (1993) and Duca and Rosenthal (1993), net worth here is instrumented in equations (3), (4), and (5).

Net worth is instrumented by variables that are correlated with net worth, but those variables have no direct effect on demand for debt, outstanding debt, or the probability of being credit constrained and are not correlated with the error terms in equations (3), (4), and (5). The instrument variables used in this study are from the literature, as follows: years of education of the spouse of the household head, whether household head or spouse received or expects to receive a large inheritance, and the natural log of expected inheritance (Cox and Jappelli 1993; Duca and Rosenthal 1993).[13] This instrumentation can proxy for a family's socioeconomic status and thus be a good candidate for proxying net worth. Also, as inheritances are concentrated among households at the top of the wealth distribution, they can differentiate households with high net worth from those with low net worth. Table 1 gives a detailed description of variables in X_1, X_2, and X_3 as well as the variables used in instrumenting net worth.

For the second step of estimating equation (3), its estimated coefficient values are used to predict the desired stock of debt for the constrained households (D^*_{1C}). In the third step, the gap between the desired and actual stock of debt is calculated for each year, as in equation (6), and the gap measure in 2007 is compared with the one in 1998.

$$GAP = 1 - \left(\frac{\overline{D_1}}{p\overline{D^*_{1U}} + (1-p)\,\overline{D^*_{1C}}} \right) \tag{6}$$

$\overline{D_1}$ is the average of the outstanding debt that is observed in the SCF data. $\overline{D^*_{1U}}$ is the average of the desired stock of debt for unconstrained households. Given that unconstrained households borrow their desired debt levels, $\overline{D^*_{1U}}$ will be equal to the average of their

[13] Lag of net worth is not used as an instrument because the data are cross-sectional.

Table 1. List of Variables and Definitions

Variable	Definition
X_1: Ln Desired Debt Level[a] or X_2: Demand for Debt	
Ln Net Worth[a]	Net worth (constructed, see appendix B) in 2007 dollars, then transformed to natural log
Homeownership	1 if owns a home
Ln Income[a]	Family's total annual income from all sources (wages, interest income, dividends, capital gains, unemployment compensation, retirement income, etc.) in 2007 dollars, then transformed to natural log
Ln Permanent Income[a]	Permanent income (constructed, see appendix C) in 2007 dollars, then transformed to natural log
Years of Education	Household head's years of education
Current Employment Status	1 if household head is unemployed and looking for work
Age	Spline function for age
Married	1 if household head is married
Black	1 if household head is black /African-American
Female	1 if household head is female
Family size	Number of people in the household
Foreseeable Large Expenses	1 if in the next 5 to 10 years, there are foreseeable large expenses (educational expenses, purchases of a new home or other durable goods, health care costs, etc.) that the household has to pay
Whether Household Feels It Is All Right to Borrow to Finance:	
Vacation Expenses	1 if yes, to finance expenses of a vacation trip
Living Expenses	1 if yes, to cover living expenses when income is cut
Purchase of Luxury Items	1 if yes, to finance the purchase of luxury items (fur coat, jewelry)
Purchase of a Car	1 if yes, to finance the purchase of a car
Educational Expenses	1 if yes, to finance educational expenses
Risk Averse	1 if household head or spouse not willing to take any financial risk when they save or make investments
Expectation about Interest Rates Five Years from Now	Household head's expectation about interest rates five years from now: 1: if higher, 2: if lower, 3: if about the same

(*continued*)

Table 1. (Continued)

Variable	Definition
X_3: Additional Variables (Has No Borrowing Constraint)	
Number of Credit Cards	Number of any type of credit cards
Has Checking Account	1 if family has checking account
Welfare	1 if family received any income from Aid to Dependent Children, Aid to Families with Dependent Children, food stamps, or other forms of welfare
Time at Current Address	How many years the family has lived within 25 miles of their current address
Time at Current Job	Number of years household head worked for the current employer
Credit History:	
Almost Always/Always Pay Off Balance on Credit Cards	1 if household always/almost always pays off total balance owed on credit cards each month
No Loan	1 if household did not have any type of loan or mortgage payments during the last year
Timely on Loan Payments	1 if during the last year all loan or mortgage payments were made as scheduled or ahead of schedule
Instrumenting Net Worth	
Spouse's Years of Education	Spouse's years of education
Whether Received Inheritance in the Past	1 if household head or spouse received an inheritance or was given substantial assets in a trust or in some other form
Expected Inheritance	1 if household head or spouse expect to receive substantial inheritance or transfer of assets
Ln Expected Inheritance[a]	Value of expected inheritance or transfer of assets (in 2007 dollars), then transformed to natural log

[a] If $x >= 0$, then $\ln(1+x)$; if $x < 0$, then $-\ln(1-x)$.

outstanding debt ($\overline{D_1}$). $\overline{D_{1C}^*}$ is the average of the desired stock of debt for constrained households, p is the probability of being unconstrained, and $(1 - p)$ is the probability of being constrained.

5. Variables

Following a large number of other studies (Cox and Jappelli 1990, 1993; Crook 1996, 2001; Duca and Rosenthal 1993; Fissell and Jappelli 1990; Gropp, Scholz, and White 1997; Jappelli 1990; Lyons 2003), this study defines households with credit constraints as those that in the preceding five years were denied credit, or could not get as much credit as they applied for, or were discouraged from applying for credit.[14] Specifically, a household is considered to have a credit constraint if it answered yes to either of the following questions:

- In the past five years, has a particular lender or creditor turned down any request you made for credit or not given you as much credit as you applied for?
- Was there any time in the past five years that you thought of applying for credit at a particular place but changed your mind because you thought you might be turned down?

Previously denied households that were later able to obtain the full loan amount they requested by reapplying to the same institution or by applying elsewhere are not included in the sample of constrained households.

Discouraged households are considered here to be constrained because the findings by Jappelli (1990) indicate that such households are similar in characteristics to those that have been turned down and could be expected to have been rejected if they had applied for credit.

The net worth variable used in estimating desired debt and the probability of having credit constraints is defined in appendix B. The effect of net worth on desired debt levels (the demand side) is

[14] This definition of being credit constrained is fairly standard in SCF literature, because SCF data does not have any information on credit scores. Although it is fairly standard, it is not a perfect measure because households may not necessarily have been credit constrained at the time of the survey, although they were at some point in the past five years.

ambiguous. On one hand, households with high net worth can have less need to borrow against their future income to smooth their consumption, and thus the amount of debt they demand can be low; on the other hand, such households have the wherewithal to demand a high level of debt desired to fund a major purchase (for instance, home purchase). On the supply side, the creditor-imposed debt ceiling varies directly with net worth—any lender would consider a household with high net worth to be less risky and would grant more or higher loans to those households. Nonetheless, the ambiguous effect of net worth on desired debt means that the effect of net worth on excess demand for debt and the probability of being unconstrained are predicted to be ambiguous.

Under the assumption that home purchases are funded with debt, homeowners are a proxy for individuals with good credit history, and the creditor-imposed debt ceiling on constrained households is therefore expected to be higher for homeowners. From the demand side, it is possible that individuals purchased their homes because they were willing to take on larger debt. Homeowners can also choose to upgrade their homes if they are willing to take on larger debt. Therefore, the effect of homeownership on desired debt is predicted to be positive. An increase in desired debt and the debt ceiling will have an ambiguous effect on the excess demand for debt and the probability of being unconstrained.

As current income increases, a household's demand for durables can increase, resulting in an increase in demand for debt. However, the increase in current income could instead reduce the need to borrow to finance current consumption and thus lower the demand for debt. On the supply side, lenders will see a household's increase in income as an increase in its ability to repay debt and will thus increase the debt ceiling and grant larger loan amounts. The ambiguous effect of higher income on the demand for debt means that it will have an ambiguous effect on the excess demand for debt and the probability of being unconstrained.

The construction of permanent income is based on Cox and Jappelli (1993) and King and Dicks-Mireaux (1981) and is detailed in appendix C. Following the life-cycle hypothesis, an increase in permanent income raises desired consumption (including of durable goods) and thus desired borrowing to finance the increase in consumption. If lenders expect a rise in the ability of households to repay their loans, they will increase the debt ceiling, resulting in an ambiguous effect on the excess demand for debt and the probability of being unconstrained.

Highly educated households are expected to have greater future income and therefore to have a higher desired consumption and demand for debt and a higher debt ceiling. Thus, the relationship between the probability of being unconstrained and education level is ambiguous. Being unemployed proxies for low expected future income, resulting in a decline in desired consumption, desired debt, and the debt ceiling. Hence, the relationship between the probability of being unconstrained and being unemployed is likewise ambiguous.

Age is defined here, as in the literature, by a spline function. According to the LCPIH, individuals have the highest level of desired stock of debt during their middle years, and a spline function for age captures the life-cycle features of desired debt. The spline function—see Cox and Jappelli (1993)—is defined as follows:

- $\text{Age}^{(1)} = \text{Age}$ if age ≤ 24; $\text{Age}^{(1)} = 24$ otherwise,
- $\text{Age}^{(2)} = \text{Min}\ (\text{Age} - 24, 10)$ if age > 24; $\text{Age}^{(2)} = 0$ otherwise,
- $\text{Age}^{(3)} = \text{Min}\ (\text{Age} - 34, 10)$ if age > 34; $\text{Age}^{(3)} = 0$ otherwise,
- $\text{Age}^{(4)} = \text{Min}\ (\text{Age} - 44, 10)$ if age > 44; $\text{Age}^{(4)} = 0$ otherwise,
- $\text{Age}^{(5)} = \text{Min}\ (\text{Age} - 54, 10)$ if age > 54; $\text{Age}^{(5)} = 0$ otherwise,
- $\text{Age}^{(6)} = \text{Age} - 64$ if age > 64; $\text{Age}^{(6)} = 0$ otherwise.

As in Crook (1996) and Jappelli (1990), age has an ambiguous relationship with the probability of being unconstrained. The debt ceiling is likely to increase with the age of the household head because the probability of default declines with the advancing age of the head. On the demand side, there are two effects. If young household heads expect their income in future periods to be much greater than their current income, they will demand more debt than older household heads. If the rate of time preference is low relative to the real rate of interest, the household will delay consumption, and the desired consumption and debt will increase with age. Because of the ambiguous effect of age on demand for debt, the net effect of age on the probability of being unconstrained is ambiguous.

As noted in Jappelli (1990), desired consumption by married couples can be lower because of economies of scale in the consumption of durables. But married couples may demand more debt, especially if one spouse is not working and is dependent on the working spouse. From the supply side, lenders will increase the debt ceiling for married couples because they are less mobile and both partners can be responsible for repaying the loan. Because of the ambiguous effect on demand for debt, excess demand for debt and the probability of being unconstrained are ambiguous for married couples.

The effect of race and gender on the probability of being unconstrained is ambiguous. Compared with others, because of their lower expected future income, nonwhites and females may have lower desired consumption and desired borrowing, and they may be granted lower amounts of credit. The result is an ambiguous effect on excess demand for loans. As family size increases, desired consumption increases and so does the desired debt level. Crook, Thomas, and Hamilton (1992) show that the probability of default rises with the number of children, which leads to a predicted decline in the debt ceiling; the result is a greater excess demand for debt and a lower probability of being unconstrained. Crook (1996) argues that households with foreseeable large expenses in the next 5 to 10 years would desire higher debt to pay for those expenses but are credit constrained. In this case, the excess demand for debt is predicted to be higher and the probability of being unconstrained is lower.

Household preferences for holding debt are proxied by whether they are risk averse. Households that are willing to borrow to finance educational and living expenses, a vacation, and purchase of luxury items and a car will hold more debt; but from the supply side, lenders will be unwilling to grant them high loan amounts. So, with the willingness to borrow, the excess amount of desired debt will increase and the probability of being unconstrained will decline. In contrast, risk-averse households will be willing to hold less debt, save more, and wait to consume at later stages in their lives when they are financially more secure. So risk-averse households are expected to desire less debt. From the supply side, the debt ceiling will be higher for risk-averse households because lenders will perceive them as being less likely to default. Given risk aversion, the excess demand for debt will decline, and the probability of being unconstrained will be greater.

Expectations about interest rates five years from now can also affect the willingness of households to consume today and borrow. If households expect interest rates to be lower in the next five years, they will probably prefer to save and invest today and wait to borrow and consume. Thus, their desired debt is expected to be low. From the supply side, lenders would prefer to grant loans at higher interest rates today, so current debt ceilings are expected to be high. With a household's expectation of lower future interest rates, the excess demand for debt is expected to decline and the probability of being unconstrained is expected to increase.

All of the above variables are included as regressors in determining both the desired stock of debt and the probability of being unconstrained. Additional regressors, which are generally requested in loan application forms, are considered in determining the probability of being unconstrained. These regressors are number of credit cards owned, whether the household has a checking account, whether the household received public assistance, the household head's time at the current job and at the current address, and whether the household had any problems making previous loan payments. Time at the current job and time at the current address proxy for the household's mobility and can be important factors in lenders' decisions on whether to grant loans.

6. Data and Descriptive Statistics

The SCF data, collected every three years under the auspices of the Board of Governors of the Federal Reserve System, contain detailed information about household balance sheets, income, use of financial services, past and current employment histories, retirement plans associated with their previous and current jobs, and demographics. The SCF sample of households is based on a dual-frame design in which the SCF oversamples wealthy households. To adjust for the oversampling, the SCF staff construct sampling weights with the post-stratification technique; the weights are provided in the data set of each survey's results.[15]

The SCF staff impute missing data with a multiple imputation technique involving six iterations. The sixth iteration produces the five implicates published in the data set—see Kennickell (1991).

[15]Kennickell, McManus, and Woodburn (1996) give detailed information on sampling and the construction of the weights.

To adjust for oversampling of wealthy households, the analysis throughout the present study takes into consideration the sampling weights. Meanwhile, the paper corrects for sampling variance by bootstrapping standard errors. The analysis is based on implicate 1 because the replicate weights for bootstrapping are valid only for implicate 1. The paper does *not* adjust for multiple imputation.[16] If statistics are not adjusted for multiple imputation and corrected for imputation uncertainty, the estimated standard error of a statistic will be smaller than the true standard error and the significance of variables will be overestimated (for instance, it may be concluded that a variable is significant at 1 percent even though it is significant at 5 percent or 10 percent, or it is not significant at all).

Observations in which permanent income is missing or family income is negative are deleted. All values are expressed in 2007 dollars. The distributions for total debt, net worth, income, permanent income, and the amount of inheritance that households expect to receive are highly skewed; a log transformation has been applied to normalize the distributions.[17]

The descriptive statistics for 1998 and 2007 were adjusted for population by using the sampling weights provided by the SCF staff. In both 1998 and 2007, the characteristics of debt-holding households and unconstrained households in comparison with other households were generally as expected (tables 2 and 3).

[16]The SCFCOMBO module (https://ideas.repec.org/c/boc/bocode/s458017.html) in Stata corrects for imputation uncertainty by estimating five sets of coefficients from five data sets, and meanwhile corrects for sampling variance by bootstrapping standard errors with the replicate weights (wgt1–wgt999) provided for implicate 1. Users can use the *scfcombo* command to run simple models; that's purely run dependent variables on several independent variables. It is not fit for advanced models such as probit models with instrumental variables and two-stage least-squares models, which are used in this paper. Therefore, in this paper, it is not possible to correct for imputation uncertainty by using the SCFCOMBO module.

It is also possible to correct for the imputation uncertainty with the user-written micombine.ado file or *mi estimate* command in Stata. For *mi estimate* and *micombine* commands, all five implicates should be used and five sets of coefficients should be estimated. However, if these commands are used, it is not possible to bootstrap standard errors, because bootstrap replication weights are valid only for the first implicate and are not available for the other (four) implicates. Therefore, in this paper, it is also not possible to correct for imputation uncertainty by using the micombine.ado file or *mi estimate* command.

[17]Following Crook and Hochguertel (2006), the following transformations have been applied: If $x < 0$, then $-\ln(1 - x)$. If $x > 0$, then $\ln(1 + x)$.

Table 2. Descriptive Statistics for 1998

	All		Debt Holders		Non-debt Holders		Difference		Unconstrained		Constrained		Difference	
	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.
Hold Debt	0.786	0.009	1.000	0.000	NA	NA	NA	NA	0.774	0.010	0.838	0.018	−0.063***	0.020
Unconstrained	0.823	0.008	0.811	0.010	0.866	0.015	−0.055***	0.018	1.000	0.000	NA	NA	NA	NA
Ln Total Debt	8.233	0.098	10.481	0.043	0.000	0.000	10.481***	0.043	8.205	0.111	8.364	0.198	−0.159	0.227
Ln Net Worth	10.684	0.127	10.627	0.153	10.895	0.181	−0.268	0.237	11.406	0.121	7.326	0.402	4.080***	0.420
Homeownership	0.722	0.009	0.766	0.010	0.562	0.023	0.204***	0.025	0.775	0.009	0.477	0.026	0.298***	0.027
Ln Income	10.728	0.032	10.890	0.033	10.138	0.078	0.752***	0.085	10.818	0.032	10.311	0.096	0.507***	0.101
Ln Permanent Income	9.462	0.033	9.807	0.028	8.200	0.089	1.607***	0.093	9.362	0.039	9.930	0.040	−0.568***	0.056
Years of Education	13.180	0.063	13.506	0.065	11.985	0.154	1.521***	0.168	13.266	0.070	12.777	0.134	0.489***	0.151
Current Employment Status	0.030	0.004	0.025	0.004	0.047	0.010	−0.022**	0.010	0.027	0.004	0.045	0.011	−0.018	0.011
Age	47.056	0.352	44.258	0.335	57.308	0.902	−13.051***	0.962	48.867	0.393	38.627	0.619	10.241***	0.733
Married	0.724	0.010	0.752	0.011	0.624	0.022	0.128***	0.025	0.742	0.010	0.641	0.024	0.101***	0.026
Black	0.085	0.006	0.078	0.007	0.109	0.014	−0.031**	0.016	0.073	0.006	0.141	0.018	−0.069***	0.019
Female	0.002	0.001	0.001	0.001	0.003	0.002	−0.002	0.002	0.001	0.001	0.004	0.003	−0.002	0.003
Family Size	2.878	0.031	3.046	0.035	2.263	0.054	0.782***	0.064	2.808	0.032	3.203	0.086	−0.395***	0.092
Foreseeable Large Expenses	0.530	0.011	0.568	0.012	0.391	0.022	0.177***	0.025	0.498	0.012	0.676	0.024	−0.177***	0.027
Attitudes Toward Credit to Finance:														
Vacation Expenses	0.137	0.007	0.154	0.009	0.074	0.011	0.080***	0.014	0.128	0.008	0.178	0.020	−0.050**	0.021
Living Expenses	0.423	0.011	0.438	0.012	0.370	0.022	0.068***	0.025	0.391	0.012	0.571	0.025	−0.180***	0.028
Purchase of Luxury Items	0.058	0.005	0.067	0.006	0.025	0.007	0.042***	0.009	0.052	0.005	0.086	0.014	−0.035**	0.015
Purchase of a Car	0.815	0.008	0.861	0.009	0.648	0.022	0.212***	0.024	0.812	0.009	0.831	0.019	−0.019	0.021
Educational Expenses	0.812	0.009	0.846	0.009	0.688	0.022	0.158***	0.024	0.797	0.010	0.880	0.017	−0.083***	0.019
Risk Averse	0.330	0.010	0.283	0.011	0.499	0.023	−0.216***	0.026	0.319	0.011	0.380	0.025	−0.061**	0.027

(*continued*)

Table 2. (Continued)

	All		Debt Holders		Non-debt Holders		Difference		Unconstrained		Constrained		Difference	
	Mean	**Std. Err.**	**Mean**	**Std. Err.**	**Mean**	**Std. Err.**	**Mean**	**Std. Err.**	**Mean**	**Std. Err.**	**Mean**	**Std. Err.**	**Mean**	**Std. Err.**
Expectation about Interest Rates Five Years from Now:														
Higher	0.661	0.010	0.681	0.011	0.589	0.023	0.091***	0.025	0.650	0.011	0.714	0.023	−0.064**	0.026
Lower	0.058	0.005	0.051	0.005	0.084	0.012	−0.033**	0.013	0.055	0.005	0.075	0.013	−0.020	0.014
About the Same	0.281	0.010	0.268	0.011	0.327	0.022	−0.058**	0.024	0.296	0.011	0.211	0.021	0.084***	0.024
Number of Credit Cards	3.879	0.084	4.343	0.096	2.179	0.153	2.164***	0.180	4.067	0.092	3.004	0.201	1.063***	0.221
Has Checking Account	0.893	0.007	0.926	0.006	0.774	0.019	0.152***	0.020	0.916	0.007	0.788	0.020	0.128***	0.021
Welfare	0.042	0.004	0.033	0.004	0.075	0.012	−0.043***	0.012	0.029	0.004	0.102	0.015	−0.073***	0.015
Time at Current Address	37.943	0.802	35.829	0.894	45.684	1.745	−9.855***	1.961	39.390	0.884	31.204	1.870	8.186***	2.069
Time at Current Job	7.291	0.199	8.101	0.224	4.323	0.403	3.778***	0.461	7.714	0.225	5.325	0.404	2.389***	0.463
Almost Always/Always Pay Off	0.419	0.011	0.397	0.012	0.499	0.023	−0.102***	0.026	0.479	0.012	0.136	0.017	0.343***	0.021
No Loan	0.206	0.009	0.000	0.000	0.961	0.008	−0.961***	0.008	0.216	0.010	0.159	0.017	0.057***	0.020
Timely on Loan Payments	0.635	0.010	0.799	0.010	0.035	0.007	0.763***	0.012	0.661	0.011	0.517	0.026	0.144***	0.028
No. of Observations	3,332													

Notes: ***, **, and * refer to significance levels at 1 percent, 5 percent, and 10 percent, respectively. The values for dummy variables are in decimals, and they need to be multiplied by 100 to be expressed in percentages.

Table 3. Descriptive Statistics for 2007

	All		Debt Holders		Non-debt Holders		Difference		Unconstrained		Constrained		Difference	
	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.
Hold Debt	0.806	0.008	1.000	0.000	NA	NA	NA	NA	0.802	0.009	0.827	0.020	−0.025	0.022
Unconstrained	0.841	0.008	0.837	0.009	0.858	0.016	−0.021	0.019	1.000	0.000	NA	NA	NA	NA
Ln Total Debt	8.746	0.096	10.853	0.044	0.000	0.000	10.853***	0.044	8.821	0.106	8.347	0.222	0.474*	0.246
Ln Net Worth	10.993	0.123	11.002	0.144	10.955	0.205	0.048	0.251	11.649	0.120	7.523	0.399	4.127***	0.417
Homeownership	0.738	0.009	0.778	0.009	0.571	0.023	0.207***	0.025	0.784	0.009	0.494	0.026	0.290***	0.028
Ln Income	10.904	0.022	11.024	0.024	10.408	0.045	0.616***	0.051	10.978	0.025	10.513	0.043	0.465***	0.049
Ln Permanent Income	10.055	0.021	10.238	0.018	9.295	0.058	0.943***	0.061	10.046	0.024	10.104	0.029	−0.058	0.037
Years of Education	13.378	0.059	13.622	0.061	12.367	0.157	1.255***	0.168	13.556	0.064	12.440	0.137	1.116***	0.152
Current Employment Status	0.023	0.003	0.017	0.003	0.048	0.010	−0.031***	0.011	0.019	0.003	0.048	0.011	−0.030**	0.012
Age	48.445	0.354	46.344	0.341	57.166	1.019	−10.822***	1.075	49.919	0.392	40.652	0.692	9.267***	0.795
Married	0.704	0.010	0.737	0.010	0.566	0.024	0.170***	0.026	0.727	0.010	0.581	0.026	0.146***	0.028
Black	0.096	0.006	0.095	0.007	0.101	0.014	−0.006	0.015	0.079	0.006	0.190	0.021	−0.112***	0.022
Female	0.005	0.002	0.005	0.002	0.007	0.004	−0.002	0.004	0.005	0.002	0.007	0.004	−0.002	0.005
Family Size	2.826	0.029	2.965	0.033	2.248	0.060	0.718***	0.068	2.774	0.031	3.099	0.085	−0.324***	0.090
Foreseeable Large Expenses	0.542	0.010	0.572	0.011	0.418	0.023	0.154***	0.026	0.522	0.011	0.647	0.025	−0.125***	0.028
Attitudes Toward Credit to Finance:														
Vacation Expenses	0.138	0.007	0.148	0.008	0.095	0.014	0.053***	0.016	0.132	0.008	0.171	0.020	−0.040*	0.021
Living Expenses	0.509	0.010	0.531	0.012	0.419	0.023	0.112***	0.026	0.479	0.011	0.671	0.025	−0.193***	0.027
Purchase of Luxury Items	0.051	0.005	0.056	0.005	0.029	0.008	0.027***	0.010	0.050	0.005	0.058	0.012	−0.009	0.013
Purchase of a Car	0.816	0.008	0.860	0.008	0.633	0.023	0.228***	0.024	0.814	0.009	0.825	0.020	−0.010	0.022
Educational Expenses	0.838	0.008	0.876	0.008	0.680	0.022	0.196***	0.024	0.826	0.009	0.899	0.016	−0.072***	0.018
Risk Averse	0.361	0.010	0.318	0.011	0.541	0.024	−0.223***	0.026	0.337	0.011	0.490	0.026	−0.153***	0.028

(*continued*)

Table 3. (Continued)

	All		Debt Holders		Non-debt Holders		Difference		Unconstrained		Constrained		Difference	
	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.
Expectation about Interest Rates Five Years from Now:														
Higher	0.656	0.010	0.657	0.011	0.653	0.023	0.005	0.025	0.644	0.011	0.723	0.024	−0.079***	0.026
Lower	0.077	0.006	0.076	0.006	0.082	0.013	−0.006	0.015	0.078	0.006	0.073	0.014	0.004	0.015
About the Same	0.267	0.009	0.267	0.010	0.266	0.021	0.001	0.024	0.279	0.010	0.204	0.021	0.075***	0.024
Number of Credit Cards	3.494	0.073	3.897	0.083	1.821	0.125	2.075***	0.150	3.773	0.080	2.019	0.160	1.753***	0.179
Has Checking Account	0.913	0.006	0.951	0.005	0.756	0.020	0.195***	0.021	0.933	0.006	0.808	0.020	0.125***	0.021
Welfare	0.065	0.005	0.053	0.005	0.119	0.015	−0.066***	0.016	0.040	0.004	0.197	0.021	−0.157***	0.021
Time at Current Address	37.049	0.748	35.959	0.827	41.570	1.729	−5.611***	1.917	37.290	0.809	35.774	1.964	1.516	2.124
Time at Current Job	7.196	0.188	7.783	0.211	4.760	0.386	3.023***	0.440	7.680	0.212	4.638	0.351	3.041***	0.410
Almost Always/Always Pay Off	0.418	0.010	0.405	0.011	0.469	0.024	−0.064**	0.026	0.472	0.011	0.131	0.018	0.341***	0.021
No Loan	0.184	0.008	0.000	0.000	0.946	0.009	−0.946***	0.009	0.186	0.009	0.171	0.020	0.016	0.022
Timely on Loan Payments	0.622	0.010	0.760	0.010	0.047	0.009	0.713***	0.013	0.652	0.011	0.464	0.026	0.188***	0.028
Number of Observations	3,511													

Notes: ***, **, and * refer to significance levels at 1 percent, 5 percent, and 10 percent, respectively. The values for dummy variables are in decimals, and they need to be multiplied by 100 to be expressed in percentages.

7. Results

Regression results for equations (3), (4), and (5) are provided for 1998 in table 4 and for 2007 in table 5. The regressions adjust for the survey nature of the SCF data by using the sampling weights, and they apply the bootstrap technique by using the replicate weights provided in the data. Table 6 displays the gaps between actual and desired debt levels. The regression results obtained from the construction of permanent income are presented and discussed in table C.2 in appendix C.

For 1998, the results for equation (3) show that homeownership, income, permanent income, and years of education had a positive relation with the desired stock of debt (table 4). A 1 percent increase in family income would increase the desired stock of debt by 0.249 percent. Likewise, a 1 percent increase in permanent income would increase the desired stock by 0.952 percent. Households between 44 and 54 years of age would desire more debt as they aged, and risk-averse households would desire less debt.

Likewise, the results for equation (3) for 2007 show a higher desired debt for homeowners and for households with high income and high permanent income, and lower desired debt for households that are risk averse in their savings or investments (table 5). The desired stock of debt was greater for household heads that were married, that had more family members, and that were willing to borrow to finance educational expenses. The unemployed demanded less debt. Younger household heads, households with foreseeable large expenses, and households that were willing to borrow to finance the purchase of a car demanded more debt, but only at the weakest (10 percent) significance level.

The effect on desired debt from variations in homeownership, income, being unemployed, family size, having foreseeable large expenses, and being risk averse are consistent with the findings of Crook (2001), Crook and Hochguertel (2006), and Duca and Rosenthal (1993). Desired debt varies significantly and directly with permanent income, a result consistent with the findings of Cox and Jappelli (1993) and Crook and Hochguertel (2006). Some of the previous literature found desired debt to be negatively related to age for household heads in the later stages of their lives (Cox and Jappelli 1993; Crook and Hochguertel 2006; Gropp, Scholz, and White 1997),

Table 4. Regression Results for 1998

	Natural Log of Stock of Debt: Equation (3)		Debt Holders: Equation (4)		Unconstrained: Equation (5)	
	Coeff.	Bootstrap Std. Err.	dy/dx	Delta-Method Std. Err.	dy/dx	Delta-Method Std. Err.
Ln Net Worth	−0.090	0.252	−0.028**	0.012	0.012	0.010
Homeownership	3.177**	1.550	0.342***	0.049	0.031	0.047
Ln Income	0.249**	0.112	0.016**	0.007	0.002	0.006
Ln Permanent Income	0.952***	0.135	0.077**	0.031	0.050	0.031
Years of Education	0.074*	0.038	0.009***	0.003	−0.002	0.003
Current Employment Status	−0.475	0.688	−0.112***	0.035	0.023	0.030
Age $\leq$ 24	0.084	0.105	0.008	0.012	−0.007	0.012
24 $<$ Age $\leq$ 34	0.031	0.052	0.005	0.006	0.001	0.004
34 $<$ Age $\leq$ 44	0.025	0.020	0.006	0.004	0.001	0.003
44 $<$ Age $\leq$ 54	0.053***	0.020	0.004	0.003	0.001	0.003
54 $<$ Age $\leq$ 64	0.004	0.109	−0.009	0.005	0.016***	0.005
64 $<$ Age	0.027	0.093	0.004	0.006	0.012**	0.006
Married	0.036	0.169	0.025*	0.015	−0.021	0.013
Black	−0.118	0.271	−0.006	0.018	−0.034	0.023
Female	−0.085	0.233	−0.080	0.117	−0.027	0.101
Family Size	0.069	0.103	0.014***	0.006	−0.009*	0.005
Foreseeable Large Expenses	−0.030	0.323	0.036***	0.012	−0.039***	0.013
Attitudes Toward Credit to Finance:						
Vacation Expenses	0.338	0.311	0.070***	0.017	−0.003	0.017
Living Expenses	−0.004	0.262	0.009	0.011	−0.040***	0.013
Purchase of Luxury Items	0.025	0.223	0.010	0.022	−0.053**	0.025
Purchase of a Car	0.299	0.408	0.087***	0.018	−0.003	0.016
Educational Expenses	−0.061	0.098	−0.025*	0.015	−0.026*	0.015
Risk Averse	−0.424**	0.192	−0.034**	0.015	−0.006	0.013

(*continued*)

Table 4. (Continued)

	Natural Log of Stock of Debt: Equation (3)		Debt Holders: Equation (4)		Unconstrained: Equation (5)	
	Coeff.	Bootstrap Std. Err.	dy/dx	Delta-Method Std. Err.	dy/dx	Delta-Method Std. Err.
Expectation about Interest Rates Five Years from Now:[a]						
Lower	−0.330*	0.180	−0.027	0.027	−0.034	0.024
About the Same	0.046	0.111	−0.011	0.014	0.008	0.015
Number of Credit Cards	NA	NA	NA	NA	0.000	0.002
Has Checking Account	NA	NA	NA	NA	0.010	0.031
Welfare	NA	NA	NA	NA	−0.013	0.028
Time at Current Address	NA	NA	NA	NA	(3.465E-04)**	1.645E-04
Time at Current Job	NA	NA	NA	NA	0.001	0.001
Almost Always/Always Pay Off	NA	NA	NA	NA	0.107***	0.024
No Loan	NA	NA	NA	NA	0.058*	0.030
Timely on Loan Payments	NA	NA	NA	NA	0.055**	0.022
Constant	−7.691*	4.306				
Inverse Mills (Debt Holders)	1.903	2.912				
Inverse Mills (Unconstrained)	0.239	2.994				
Number of Observations	2,155		3,332		3,332	
Prob > chi2	0.0000		0.0000		0.0000	
athrho			0.501*	0.304	−0.228	0.305

Note: ***, **, and * refer to significance levels at 1 percent, 5 percent, and 10 percent, respectively.
[a]Omitted category in expectation about interest rates five years from now is "higher."

Table 5. Regression Results for 2007

	Natural Log of Stock of Debt: Equation (3)		Debt Holders: Equation (4)		Unconstrained: Equation (5)	
	Coeff.	Bootstrap Std. Err.	dy/dx	Delta-Method Std. Err.	dy/dx	Delta-Method Std. Err.
Ln Net Worth	−0.622	0.398	−0.046***	0.009	−0.026**	0.011
Homeownership	8.944**	4.199	0.404***	0.028	0.180***	0.058
Ln Income	0.620*	0.328	0.027***	0.009	−0.002	0.010
Ln Permanent Income	2.288**	0.983	0.085***	0.021	0.108***	0.019
Years of Education	0.015	0.028	0.004	0.003	0.000	0.003
Current Employment Status	−3.200*	1.904	−0.184***	0.052	−0.047	0.038
Age $\leq$ 24	0.714*	0.419	0.030*	0.015	0.008	0.012
24 $<$ Age $\leq$ 34	−0.019	0.033	3.11E-05	0.004	0.001	0.003
34 $<$ Age $\leq$ 44	0.101	0.069	0.005*	0.003	0.001	0.002
44 $<$ Age $\leq$ 54	0.107	0.080	0.004	0.003	0.013***	0.003
54 $<$ Age $\leq$ 64	0.011	0.017	−0.001	0.003	0.003	0.003
64 $<$ Age	−0.021	0.023	−0.002	0.003	0.018***	0.003
Married	1.007**	0.413	0.047***	0.016	−0.003	0.017
Black	−0.319	0.415	−0.004	0.024	−0.060***	0.022
Female	−0.231	0.742	−0.046	0.074	0.138***	0.047
Family Size	0.143**	0.063	0.014***	0.005	0.005	0.005
Foreseeable Large Expenses	0.446*	0.248	0.035***	0.012	−0.029**	0.012
Attitudes Toward Credit to Finance:						
Vacation Expenses	0.042	0.230	0.000	0.018	−0.021	0.017
Living Expenses	0.086	0.094	0.012	0.011	−0.037***	0.011
Purchase of Luxury Items	−0.118	0.194	0.005	0.026	−0.026	0.027
Purchase of a Car	1.637*	0.885	0.112***	0.017	−0.005	0.016
Educational Expenses	0.322**	0.128	0.018	0.015	−0.053***	0.017
Risk Averse	−1.005**	0.453	−0.060***	0.013	−0.036***	0.014

(*continued*)

Table 5. (Continued)

	Natural Log of Stock of Debt: Equation (3)		Debt Holders: Equation (4)		Unconstrained: Equation (5)	
	Coeff.	Bootstrap Std. Err.	dy/dx	Delta-Method Std. Err.	dy/dx	Delta-Method Std. Err.
Expectation About Interest Rates Five Years from Now[a]:						
Lower	−0.066	0.185	−0.003	0.019	0.013	0.019
About the Same	−0.026	0.094	0.009	0.011	0.005	0.014
Number of Credit Cards	NA	NA	NA	NA	0.005**	0.002
Has Checking Account	NA	NA	NA	NA	0.041	0.025
Welfare	NA	NA	NA	NA	−0.181***	0.041
Time at Current Address	NA	NA	NA	NA	0.000	0.000
Time at Current Job	NA	NA	NA	NA	0.003***	0.001
Almost Always/Always Pay Off	NA	NA	NA	NA	0.125***	0.020
No Loan	NA	NA	NA	NA	0.114***	0.025
Timely on Loan Payments	NA	NA	NA	NA	0.076***	0.019
Constant	−44.699*	25.894	NA	NA	NA	NA
Inverse Mills (Debt Holders)	8.495*	4.931				
Inverse Mills (Unconstrained)	1.802	1.751				
Number of Observations	2,350		3,511		3,511	
Prob > chi2	0.0000		0.0000		0.0000	
athrho			0.916***	0.247	0.682***	0.256

Note: ***, **, and * refer to significance levels at 1 percent, 5 percent, and 10 percent, respectively.
[a]Omitted category in expectation about interest rates five years from now is “higher.”

whereas this study finds a positive relation for household heads that are 44 to 54 (as well as, like other studies, in their 20s).

For 1998, the results for equation (4) show that net worth had a negative relation with the likelihood of holding debt (table 4). For instance, if net worth increased from \$5,000 to \$15,000, the probability of holding debt declined by about 3 percent.[18] Households that owned homes were more likely to hold debt, by 34 percent, than other households. Income, permanent income, and years of education also had a positive effect on the probability of holding debt. Unemployed and risk-averse households were less likely to hold debt. More likely to hold ddebt were households with large families, or with foreseeable large expenses that they had to pay in the next 5 to 10 years, or that were willing to borrow to finance their vacation expenses and car purchases.

The findings for likelihood of holding debt in 2007 (table 5) are mostly similar to those for 1998. The notable differences are that, in 2007, household heads that were younger than 24 or between 34 and 45 were more likely to hold debt as they aged, but only at the 10 percent significance level. Also, in 2007 unlike 1998, education level and attitudes toward credit to finance vacations or educational expenses did not have any effect on the probability of holding debt.

With respect to the probability of being unconstrained, the results for equation (5) show that, in 1998, household heads aged 55 years or older were more likely to be unconstrained as they aged (table 4). Also more likely to be unconstrained were households that stayed longer in their current neighborhood, had a good credit history, almost always or always paid off their total balances in their credit cards, and were timely on their loan payments. Less likely to be unconstrained were households with large foreseeable expenses in the next 5 to 10 years and households that were willing to borrow to finance their living expenses, purchases of luxury items, and educational expenses.

The findings for 2007 show that net worth had a negative effect on the probability of being unconstrained (table 5). Households that owned homes were more likely to be unconstrained (by 18 percent), as were households with more permanent income, more credit cards,

[18] That is, (ln \$15,000 – ln \$5,000) * 0.028 = 0.0308.

no welfare payments, good job tenure, and a good record of credit card and loan payments. Households were more likely to be unconstrained as they aged, particularly those in the group between ages 44 and 55 and older than 64. Households were less likely to be unconstrained if the household head was black and male, or if the households had foreseeable large expenses in the next 5 to 10 years, were willing to borrow to finance living and educational expenses, and were risk averse in their savings or investments. These results on probability of being unconstrained are generally consistent with the findings of Cox and Jappelli (1993), Crook (1996, 2001), Duca and Rosenthal (1993), and Jappelli (1990).

Table 6 shows averages of desired and actual stocks of debt for age and income subgroups of unconstrained and constrained households in 1998 and 2007, along with the gap between the two debt measures. It indicates that (i) households overall had higher desired debt levels and lower probabilities of being credit constrained in 2007 than in 1998, (ii) each subgroup of constrained and unconstrained households desired higher levels of debt in 2007, and (iii) only some of the subgroups had a lower probability of being credit constrained in 2007 than in 1998 (34 or younger, or between 44 and 65, or had income greater than $60,000).[19]

Overall, the borrowing gap increased from about 38.91 percent in 1998 to about 41.60 percent in 2007 (table 6). But the gap declined for some households, namely those with household heads 34 or younger, or between 54 and 65, or with income between $30,000

[19] When the analysis is extended to periods 2001 and 2004, the results show an evolving trend for decline in the probability of being credit constrained.

From 2004 to 2007, for the overall sample (without classifying households into age or income groups), there was a decline in the probability of households being credit constrained. This result implies that a lower percentage of households were denied for credit.

The decline in the probability is also observed as follows:

- from 1998 to 2001, for households in age groups ($24 < \text{age} \leq 34$) and ($44 < \text{age} \leq 54$) and for households with income > $60,000;
- from 2001 to 2004, for households in age groups ($\text{age} \leq 24$) and ($\text{age} > 64$) and for households with income > $100,000;
- from 2004 to 2007, for households in age group ($\text{age} \leq 64$) and for households with income ≤ $345,000.

Table 6. Gap Calculations

	1998					
	Unconstrained	**Constrained**		**Probability**		**Average Outstd. Debt**
	Desired Debt = Actual Debt	**Actual Debt**	**Desired Debt**	**Unconstrained**	**Constrained**	
All	\$3,661	\$4,291	\$16,666	0.8076	0.1924	\$3,765
Subgroups						
Age ≤ 24	\$1,846	\$151	\$4,640	0.6663	0.3337	\$876
24 < Age ≤ 34	\$11,405	\$2,743	\$15,007	0.6692	0.3308	\$7,462
34 < Age ≤ 44	\$19,137	\$7,073	\$18,965	0.7695	0.2305	\$15,707
44 < Age ≤ 54	\$16,860	\$10,105	\$27,620	0.8072	0.1928	\$15,510
54 < Age ≤ 64	\$3,528	\$9,861	\$19,953	0.8863	0.1137	\$3,926
64 < Age	\$45	\$4,166	\$7,303	0.9650	0.0350	\$52
Income ≤ 30,000	\$153	\$365	\$6,204	0.7493	0.2507	\$190
30,000 < Income ≤ 60,000	\$2,125	\$8,074	\$20,283	0.7876	0.2124	\$2,813
60,000 < Income ≤ 100,000	\$16,805	\$29,549	\$34,962	0.8529	0.1471	\$18,285
100,000 < Income ≤ 345,000	\$29,500	\$56,686	\$58,605	0.9156	0.0844	\$31,032
345,000 < Income	\$22,022	\$17,353	\$165,981	0.9530	0.0470	\$21,766

(*continued*)

Table 6. (Continued)

	2007							
	Unconstrained	**Constrained**		**Probability**		**Average Outstd. Debt**	**GAP**	
	Desired Debt = Actual Debt	**Actual Debt**	**Desired Debt**	**Unconstrained**	**Constrained**		**1998**	**2007**
All	\$6,778	\$4,218	\$28,732	0.8185	0.1815	\$6,285	0.3891	0.4160
Subgroups								
Age ≤ 24	\$2,259	\$137	\$4,907	0.7253	0.2747	\$1,197	0.6847	0.5993
24 < Age ≤ 34	\$19,957	\$3,605	\$24,619	0.7325	0.2675	\$13,326	0.4076	0.3716
34 < Age ≤ 44	\$25,541	\$8,579	\$43,816	0.7360	0.2640	\$20,198	0.1775	0.3348
44 < Age ≤ 54	\$20,655	\$10,083	\$55,577	0.8180	0.1820	\$18,536	0.1808	0.3138
54 < Age ≤ 64	\$9,291	\$2,737	\$22,098	0.8908	0.1092	\$8,397	0.2724	0.2145
64 < Age	\$232	\$1,281	\$26,439	0.9476	0.0524	\$255	0.8275	0.8413
Income ≤ 30,000	\$197	\$381	\$11,013	0.7380	0.2620	\$232	0.8865	0.9236
30,000 < Income ≤ 60,000	\$3,858	\$6,290	\$23,560	0.7779	0.2221	\$4,294	0.5298	0.4785
60,000 < Income ≤ 100,000	\$26,873	\$59,689	\$153,275	0.8784	0.1216	\$29,645	0.0612	0.2983
100,000 < Income ≤ 345,000	\$52,107	\$122,725	\$139,901	0.9545	0.0455	\$54,232	0.0289	0.0333
345,000 < Income	\$26,557	\$6,359	\$178,955	0.9711	0.0289	\$25,469	0.2438	0.1772

Note: The average dollar values are computed from the averages of natural log values.

and \$60,000 or above \$345,000.[20] The decline in the gap implies that households on average obtained a greater fraction of their desired debt. For instance, among the younger age groups in 2007, the group 24 or younger borrowed more and obtained about 9 percentage points more of their desired debt, and the group between 54 and 65 obtained about 6 percentage points more of their desired debt. The group with income between \$30,000 and \$60,000 obtained about 5 percentage points more of their desired debt in 2007.[21]

[20] The decline in the gap for the group with income above \$345,000 is consistent with earlier findings that high-income households may also have benefited from the relaxed lending standards (Foote et al. 2008; Keys et al. 2009).

[21] The analysis is extended to periods 2001 and 2004.

From 2001 to 2004 and 2004 to 2007, for the overall sample (without classifying households into age or income groups), there was an increase in desired debt for both unconstrained and constrained households. Meanwhile, the gap declined and credit-constrained households obtained a greater fraction of their desired debt.

From 1998 to 2001, for households with income $>$ \$345,000, there was a decline in desired debt for both unconstrained and constrained households in this income group. Meanwhile, the gap declined.

From 2001 to 2004:

- For any age group, there was an increase in desired debt for unconstrained households and also for constrained households (except for the constrained households in age group ($24 < \text{age} \leq 34$).
- For households in age groups ($\text{age} \leq 24$) and ($44 < \text{age} \leq 54$) and for households with \$30,000 $<$ income $\leq$ \$100,000, there was an increase in desired debt for unconstrained and constrained households in this age and income group. Meanwhile, the gap declined and credit-constrained households obtained a greater fraction of their desired debt.

From 2004 to 2007:

- There was an increase in desired debt for unconstrained households in age groups ($24 < \text{age} \leq 34$ and $\text{age} > 54$) and constrained households in age groups ($24 < \text{age} \leq 54$ and $\text{age} > 64$). For any age group, the gap declined and credit-constrained households obtained a greater fraction of their desired debt.
- There was an increase in desired debt for unconstrained households with \$60,000 $<$ income $\leq$ \$345,000 and for constrained households with income $\leq$ \$345,000. For households with income $\leq$ \$30,000 and income $>$ \$100,000, the gap declined and credit-constrained households obtained a greater fraction of their desired debt.

8. Conclusions

This study is motivated by the 2007 subprime crisis and the findings in the literature that declines in lending standards and increases in household debt were some of the contributors to the crisis. The goal here is to examine changes in household borrowing behaviors and debt following the decline in lending standards. Using the 1998 and 2007 data collected under the sponsorship of the Board of Governors of the Federal Reserve System in its triennial Survey of Consumer Finances, this study aims to determine whether the decline in lending standards in the years between those two surveys was accompanied by (i) a desire by households for more debt, (ii) a decrease in the likelihood that households would be credit constrained (that is, an increase in the likelihood that they could get all the credit they desired), and (iii) an increase in the proportion of desired debt obtained by credit-constrained households.

The analysis uses the self-reported credit-constraint indicators from the survey data in computing the gap between households' desired and outstanding debt levels. It compares the gap in 1998 (when subprime lending was in effect but before the decline in lending standards) with the gap in 2007 (when lending standards had been in effect for some time). Desired debt has long been a topic of study, but this is the first analysis to look at households' desired debt in the run-up to the subprime mortgage crisis.

Because desired debt is not directly observed, the gap measure is constructed in three steps, as in Cox and Jappelli (1993). The first step is to estimate the desired debt for indebted and unconstrained households. The second step is to compute the desired debt for constrained households by using the estimated coefficients from the first step. The third step is to calculate the difference between the predicted debt and the observed debt levels. This paper applies these steps for 1998 and 2007.

The paper also extends the literature on household debt by including the 1998 and 2007 SCF data on wealthy households. (Those households are oversampled by the SCF; the present analysis corrects for the oversampling by using the sampling weights that are provided in the SCF data.) This paper also instruments net worth, constructs a permanent-income variable, controls the regressions for the constructed permanent-income variable, and bootstraps standard errors with the replicate weights that are provided in the data.

The findings show that (i) households desired more debt in 2007 than in 1998; (ii) the likelihood of being credit constrained declined from 1998 to 2007 for households headed by someone who was 34 years of age or younger or between the ages of 44 and 65, and for households with an annual income above $60,000; and (iii) credit-constrained households received more of the credit they desired in 2007 than in 1998 if they were headed by someone 34 years of age or younger or between the ages of 54 and 65, or if they had an income between $30,000 and $60,000 or more than $345,000.

These results suggest that it is well worth the effort for regulators to track these attributes of borrowing behaviors and debt as part of their intensified efforts to preserve financial stability.

Appendix A

$$Max_{C_1} F\left(C_1\right) = U(C_1) + \mathrm{U}[(1+r)\left(A_1 + Y_1 - C_1\right) + Y_2] \tag{A.1}$$

subject to

$$C_1 \le A_1 + Y_1 + \min\left(BC, \frac{y_2}{1+r}\right) \tag{A.2}$$

$$C_1 \ge 0 \tag{A.3}$$

$$\begin{gathered}
F'(C_1) = C_1^{-\alpha} + \left(Y_2 + (1+r)\left(A_1 + Y_1 - C_1\right)\right)^{-\alpha}\left(-\left(1+r\right)\right) = 0 \\
C_1^{-\alpha} = (1+r)\left(Y_2 + (1+r)\left(A_1 + Y_1 - C_1\right)\right)^{-\alpha} \\
C_1 = (1+r)^{-1/\alpha}\left(Y_2 + (1+r)\left(A_1 + Y_1 - C_1\right)\right) \\
C_1 = (1+r)^{-1/\alpha}\left(Y_2 + (1+r)\left(A_1 + Y_1\right)\right) - (1+r)^{1-\frac{1}{\alpha}} C_1 \\
C_1^* = \frac{(1+r)^{-1/\alpha}\left(Y_2 + (1+r)\left(A_1 + Y_1\right)\right)}{1 + (1+r)^{1-\frac{1}{\alpha}}}
\end{gathered}$$

In order to check whether the critical point C_1^* is a maximum point or not, $F\left(C_1\right)$is written as $U(C_1) + \mathrm{U}\left(\mathrm{R} - \beta C_1\right)$, where

$R = Y_2 + (1 + r)(A_1 + Y_1)$ and $\beta = (1 + r)$.

$$F'(C_1) = C_1^{-\alpha} + (R - \beta C_1)^{-\alpha}(-\beta)$$

$F''(C_1) = -\alpha C_1^{-\alpha-1} + \beta^2(-\alpha)(R - \beta C_1)^{-\alpha-1} < 0$. Thus, the critical point is a maximum point.

The range that C_1 can take is $(0, \mathrm{A}_1 + Y_1 + \min(BC, \frac{Y_2}{1+r}))$.

Denote $\mathrm{A}_1 + Y_1 + \min(BC, \frac{Y_2}{1+r})$ as Q.

Let P be the optimal solution for C_1.

If $P \leq \mathrm{Q}$, there is no boundary solution. Thus, $C_1^* = P = \frac{(1+r)^{-1/\alpha}(Y_2+(1+r)(A_1+Y_1))}{1+(1+r)^{1-\frac{1}{\alpha}}}$.

If $P > \mathrm{Q}$, there is a boundary solution. Thus, $C_1^* = \mathrm{Q} = \mathrm{A}_1 + Y_1 + \min\left(BC, \frac{Y_2}{1+r}\right)$.

Appendix B

B.1 "Total Loan" Derivation

The dependent variable "total loan" is the outstanding debt level in credit or store cards, mortgage loans on principal residence, other loans received for the purchase of principal residence, home improvement loans, loans received for the purchase of investment real estate and vacation properties, lines of credit, vehicle loans, education loans, and other consumer loans. All the outstanding debt levels are expressed in 2007 dollars.

B.2 "Net Worth" Derivation

Net Worth = Financial Assets + Nonfinancial Assets – Debt

Financial assets include dollar value accumulated in pension accounts and IRA/Keogh accounts, savings/money market accounts, annuities, trusts, and managed investment accounts, checking accounts, certificates of deposit, mutual funds, savings bonds, bonds other than savings bonds, publicly traded stocks, call money accounts/cash at stock brokerages, and cash-value insurance.

Nonfinancial assets include value of business interests, owned vehicles, equity in principal residence, real estate investments, vacation properties, and any other assets (artwork, precious metals, antiques, oil and gas leases, futures contracts, future proceeds from a lawsuit or estate, etc.).

Debt includes outstanding loans on pension plans from current job, margin loans at a stock brokerage, outstanding loans on cash-value insurance policies, outstanding vehicle loans, outstanding loans or mortgages on investment real estate, vacation properties, and principal residence; any other outstanding loans used to purchase principal residence; outstanding loans on principal residence improvement; outstanding loans on credit/store card debts; lines of credit; any money owed to business; education loans; other consumer loans; and any other debt not recorded earlier.

Appendix C

Based on King and Dicks-Mireaux (1981),

$$\log Y_i = \gamma Z_i - C(A_i) + s_i, \tag{C.1}$$

where Y_i is the permanent income of individual $i; Z_i$ is a vector of observable variables (education and occupation) for individual $i; \gamma$ is the associated coefficient for $Z_i; A_i$ is individual i's age in the sample year; $C(A_i)$ is a cohort effect, which reflects the fact that for a given Z, younger generations are better off than older generations due to technological advances and capital accumulation; and s_i is an error term that measures individual i's unobservable characteristics (skill, good fortune, etc.). Also, s_i has a population mean of zero and a variance of σ_s^2.

$$\log E_{it} = \log Y_i + h(A_{it} - \overline{A}) + u_{it}, \tag{C.2}$$

where E_{it} is the earnings for individual i at period $t, \overline{A}$ is a standard age with respect to which permanent income is defined, $h(A_{it} - \overline{A})$ is the age earnings profile which is assumed to be constant across the population, and u_{it} is the unobservable transitory component of earnings.[22] Moreover, u_{it} is uncorrelated with s_i and has a mean of zero and a variance of σ_u^2.

[22]The age earnings profile is the mean earnings of workers at various ages.

Combining equations (C.1) and (C.2) gives earnings equation (C.3).

$$\log E_{it} = \gamma Z_i + g\,(A_{it}) + s_i + u_{it}, \tag{C.3}$$

where $g\,(A_{it}) = h\,(A_{it} - \overline{A}) - C\,(A_{it})$ and is approximated by a cubic function of age. Additionally, $s_i + u_{it}$ is the error term and has a mean of zero and a variance of $\sigma_s^2 + \sigma_u^2$.

There are two ways to estimate permanent income. The first estimate is based on equation (C.1). Using equation (C.1) and omitting the unobservable individual effects s_i in equation (C.1), permanent income can be estimated as in equation (C.4).

$$\log\, Y_i^e = \gamma Z_i - C\,(A_{it}) \tag{C.4}$$

The second estimate is based on equation (C.2). Permanent income is estimated as in equation (C.5), by using the age earnings profile and the information contained in the observation of current earnings. This second estimate includes the unobservable transitory component of earnings.

$$\log\, Y_i^e = \log\, E_{it} - h\,(A_{it} - \overline{A}) \tag{C.5}$$

Following King and Dicks-Mireaux (1981), a more efficient estimate is obtained by taking a weighted average of (C.4) and (C.5). This procedure gives (C.6):

$$\log\, Y_i^e = \alpha\,\{\log\, E_{it} - h\,(A_{it} - \overline{A}\,)\} + (1-\alpha)\,\{\gamma Z_i - C\,(A_{it})\}. \tag{C.6}$$

Substituting (C.3) into (C.6) gives (C.7):

$$\log\, Y_i^e = \gamma Z_i - C\,(A_{it}) + \alpha\,(s_i + u_{it}). \tag{C.7}$$

Based on Cox and Jappelli (1993) and King and Dicks-Mireaux (1981), permanent income is predicted separately for men and women in each year, 1998 and 2007. The category "men" includes household heads that are male or spouses of female household heads. The category "women" includes household heads that are female or spouses of male household heads.[23] The steps below describe in

[23]Household heads with a spouse from the same gender are excluded. In 2007 SCF (including all the five implicates), out of 17,560 male household heads, 50 of

detail the estimation of permanent income separately for men and women in 2007. The same steps are applied for 1998. While predicting permanent income, individuals with negative earnings are deleted from each of the samples.[24]

As the first step, earnings equation (C.3) is estimated separately for each of the following two samples: men with annual earnings greater than \$8,000 and women with annual earnings greater than \$8,000.[25] Individuals with zero earnings or annual earnings less than or equal to \$8,000 are excluded from the estimation.[26] Any possible bias that may arise from excluding individuals with low or zero earning levels is corrected by Heckman MLE (maximum-likelihood estimation). In the first stage, a selection equation is estimated for the full sample to predict individuals with annual earnings greater than \$8,000 (table C.1, panel A). The selection equation controls for determinants of zero or low earnings. The Heckman MLE results are displayed in table C.2.

As the second step, permanent income is predicted separately for the following samples: men with annual earnings greater than \$8,000 ($\$8{,}000 < E_i$); men with annual earnings less than or equal to \$8,000 ($\$0 \leq E_i \leq \$8{,}000$); women with annual earnings greater

them have a spouse from the same gender. Out of 4,530 female household heads, 15 of them have a spouse from the same gender.

In 1998 SCF (including all the five implicates), out of 16,805 male household heads, 165 of them have a spouse from the same gender. Out of 4,720 female household heads, 115 of them have a spouse from the same gender.

[24]The "men" sample, in the 2007 SCF, has a total of 17,575 observations, and 20 of those observations have negative earnings. Likewise, the "women" sample has a total of 19,320 observations, and 15 of those observations have negative earnings. In the 1998 SCF, the men sample has a total of 16,670 observations, and 10 of those observations have negative earnings. The women sample has a total of 18,355 observations, and 5 of those observations have negative earnings.

[25]King and Dicks-Mireaux (1981) use \$2,000 as the earnings level. This earnings level corresponds to the 1976 data. The \$2,000 in 1976 is equivalent to \$8,000 in 2007 and \$6,000 in 1998. Therefore, this paper uses an earnings level of \$8,000 for the 2007 survey and \$6,000 for the 1998 survey, while estimating earnings equations for individuals with earnings above a certain level.

[26]Following King and Dicks-Mireaux (1981), earnings equation (12) assumes that individuals are in full-time employment. Hence the earnings equation is estimated for individuals with earnings levels greater than \$6,000 (\$8,000). For a detailed argument in favor of excluding individuals with low earnings, please refer to King and Dicks-Mireaux (1981).

Table C.1. Estimation of Permanent Income

Variables	**Definitions**
A. Selection Equation	
Dependent Variable	1 if annual earnings greater than $6,000 in the 1998 survey or $8,000 in the 2007 survey, 0 if less than or equal to $6,000 ($8,000) or equal to zero.
Marital Status	1 if married
Education	1 if elementary or less education
Age < 22	1 if age less than 22 years old
Age > 65	1 if age greater than 65 years old
Nonworker	1 if not currently employed/not self-employed or waiting to start work
Part-Time Worker	1 if current job is part time
Financial Assets	Financial assets for the family (self-constructed, see appendix B)
Number of Children under Age 18	Number of children under 18 that live in the housing unit
B. Earnings Equation	
Dependent Variable	Natural log of annual earnings
Age	Age
Agesq	Age squared
Age-cubed	Age cubed
Education	1 if none or elementary school 2 if middle school or some high school 3 if high school graduate 4 if some college 5 if college graduate 6 if graduate or professional school
Occupation	0: not doing any work for pay 1: manager and professional 2: technical, sales, and administrative 3: services 4: production, craft, and repair 5: operators and laborers 6: farming, forestry, and fishery
Race	1 if household head is black
Marital Status	1 if married

Note: While calculating annual earnings, it is assumed that there are 52 weeks in a year, and for hourly workers it is assumed that they work 40 hours a week.

Table C.2. Regression Results for Permanent Income

	Probability of Having Earnings Greater than $6,000				Probability of Having Earnings Greater than $8,000			
	1998				2007			
	Female		Male		Female		Male	
	Coeff.	Linearized Std. Err.	Coeff.	Linearized Std. Err.	Coeff.	Linearized Std. Err.	Coeff.	Linearized Std. Err.
Marital Status	−0.328***	0.091	−0.109	0.102	−0.175*	0.091	0.027	0.099
Education	−0.422***	0.102	−0.486*	0.266	−0.500***	0.170	−0.083	0.371
Age < 22	0.386	0.324	0.067	0.219	0.358	0.350	−0.294	0.304
Age > 65	−0.639***	0.155	−0.801***	0.139	−0.694***	0.126	−0.493***	0.139
Nonworker	−3.108***	0.094	−2.761***	0.102	−3.259***	0.093	−2.994***	0.100
Part-Time Worker	−0.140	0.168	0.146	0.243	−0.357**	0.172	0.278	0.280
Financial Assets	−2.06E-08	1.97E-08	(7.08E-08)***	1.80E-08	−7.07E-09	1.01E-08	(2.39E-08)**	9.96E-09
Number of Children under Age 18	−0.055	0.039	0.126***	0.045	−0.030	0.033	0.033	0.047
Constant	1.896***	0.088	1.744***	0.095	1.946***	0.091	1.747***	0.090

(*continued*)

Table C.2. (Continued)

	Natural Log of Annual Earnings							
	1998				2007			
	Female		Male		Female		Male	
Age	0.081***	0.023	0.116***	0.027	0.082***	0.029	0.080***	0.025
Agesq	−0.001***	0.001	−0.002***	0.001	−0.001*	0.001	−0.001*	0.001
Age-cubed	(6.92E-06)*	3.75E-06	(8.51E-06)*	4.51E-06	4.80E-06	4.96E-06	3.71E-06	3.87E-06
Education:								
Middle or Some High School	0.044	0.092	0.173*	0.090	0.034	0.107	0.209***	0.068
High School Graduate	0.182**	0.081	0.301***	0.085	0.196*	0.100	0.462***	0.059
Some College	0.296***	0.082	0.414***	0.085	0.389***	0.102	0.522***	0.063
College Graduate	0.422***	0.087	0.561***	0.092	0.618***	0.103	0.840***	0.068
Graduate or Professional School	0.645***	0.090	0.768***	0.098	0.903***	0.110	0.986***	0.077
Occupation:								
Manager and Professional	0.467***	0.116	0.384***	0.118	0.294**	0.127	0.660***	0.087
Technical, Sales, and Administrative	0.223*	0.115	0.174	0.118	0.168	0.128	0.342***	0.088
Services	−0.005	0.118	−0.156	0.118	−0.055	0.130	0.186**	0.088
Production, Craft, and Repair	0.346***	0.128	0.093	0.113	0.109	0.152	0.308***	0.083
Operators and Laborers	0.132	0.121	0.019	0.113	−0.019	0.138	0.242***	0.086

(*continued*)

Table C.2. (Continued)

	Natural Log of Annual Earnings							
	1998				2007			
	Female		Male		Female		Male	
Race	0.009	0.044	−0.156***	0.043	−0.062	0.040	−0.163***	0.048
Marital Status	−0.050	0.030	0.131***	0.030	−0.028	0.029	0.142***	0.031
Constant	8.085***	0.338	7.691***	0.388	8.240***	0.410	8.024***	0.361
athrho	−0.177***	0.061	−0.374***	0.056	−0.253***	0.067	−0.274***	0.057
lnsigma	−0.654***	0.031	−0.501***	0.022	−0.597***	0.020	−0.462***	0.019
rho	−0.175	0.059	−0.357	0.049	−0.248	0.063	−0.267	0.053
sigma	0.520	0.016	0.606	0.013	0.550	0.011	0.630	0.012
lambda	−0.091	0.031	−0.216	0.031	−0.136	0.035	−0.168	0.034
Number of Observations	3,660		3,326		3,838		3,498	
Prob > F	0.0000		0.0000		0.0000		0.0000	

Notes: ***, **, and * refer to significance levels at 1 percent, 5 percent, and 10 percent, respectively. Omitted category in education is "none or elementary school." Omitted category in occupation is "farming, forestry, and fishery."

than \$8,000 (\$8,000 $<E_i$); and women with annual earnings less than or equal to \$8,000 (\$0 $\leq E_i \leq$ \$8,000) (table C.1, panel B).

In order to predict permanent income for men with annual earnings greater than \$8,000, equation (C.7) is estimated. While estimating equation (C.7), the same vector of explanatory variables (Z), estimated γ coefficient values, and estimated residuals from step 1, earnings equation (C.3), are used. Standard age is taken as 45 and the cohort effect is taken as 0.75 percent. For instance, $C\,(A_i)$= (Age-45)*(0.0075). The α value is assumed to be 0.5.

However, for men with \$0 $\leq E_i \leq$ \$8,000, permanent income is predicted by equation (C.4), but by using the same vector of explanatory variables (Z) and the estimated γ coefficient values from earnings equation (C.3) in step 1. The cohort effect is still taken as 0.75 percent and standard age is taken as 45.

An initial permanent income is estimated for women with \$0 $\leq E_i \leq$ \$8,000 or \$8,000 $< E_i$ by applying the same procedures as used for males. In contrast to the procedure for males, a final permanent income is estimated by using the initial permanent-income estimate and adjusting it for the possibility of nonparticipation in the labor force by women and using the estimated probabilities of having low or no earnings.

$$Y_i^w = Y_i^e\; prob\,(E_i > \$8,000) + \overline{E_w}\; prob\,(\$0 \leq E_i \leq \$8,000)$$

Y_i^e is the initial permanent-income estimate for women, E_i is annual earnings in the sample year, and $\overline{E_w}$ is mean earnings of women with annual earnings less than or equal to \$8,000 (\$0 $\leq E_i \leq$ \$8,000). The probabilities of earnings being above and below \$8,000 are computed for each woman in the sample from the probit regression estimated in the first stage of the Heckman model in step 1.

Permanent income for households is constructed by summing the computed permanent incomes for household head and the spouse.

References

Altonji, J. G., and A. Siow. 1987. "Testing the Response of Consumption to Income Changes with (Noisy) Panel Data." *Quarterly Journal of Economics* 102 (2): 293–328.

Attanasio, O. P., and G. Weber. 1995. "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *Journal of Political Economy* 103 (6): 1121–57.

Cox, D., and T. Jappelli. 1990. "Credit Rationing and Private Transfers: Evidence from Survey Data." *Review of Economics and Statistics* 72 (3): 445–54.

———. 1993. "The Effect of Borrowing Constraints on Consumer Liabilities." *Journal of Money, Credit and Banking* 25 (2): 197–213.

Crook, J. N. 1996. "Credit Constraints and US Households." *Applied Financial Economics* 6 (6): 477–85.

———. 2001. "The Demand for Household Debt in the USA: Evidence from the 1995 Survey of Consumer Finance." *Applied Financial Economics* 11 (1): 83–91.

———. 2005. "The Measurement of Household Liabilities: Conceptual Issues and Practice." Working Paper No. 01/05, Credit Research Centre, University of Edinburgh.

Crook, J., and S. Hochguertel. 2006. "Household Debt and Credit Constraints: Comparative Micro Evidence from Four OECD Countries." Working Paper, University of Edinburgh (June).

Crook, J. N., L. C. Thomas, and R. Hamilton. 1992. "The Degradation of the Scorecard over the Business Cycle." *IMA Journal of Mathematics Applied in Business and Industry* 4: 111–23.

Dejuan, J. P., and J. J. Seater. 1999. "The Permanent Income Hypothesis: Evidence from the Consumer Expenditure Survey." *Journal of Monetary Economics* 43 (2): 351–76.

Dell'Ariccia, G. D., D. Igan, and L. Laeven. 2008. "Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market." IMF Working Paper No. 08/106.

Demyanyk, Y., and O. Van Hemert. 2011. "Understanding the Subprime Mortgage Crisis." *Review of Financial Studies* 24 (6): 1848–80.

Duca, J. V., and S. S. Rosenthal. 1993. "Borrowing Constraints, Household Debt, and Racial Discrimination in Loan Markets." *Journal of Financial Intermediation* 3 (1): 77–103.

Fissell, G. S., and T. Jappelli. 1990. "Do Liquidity Constraints Vary over Time? Evidence from Survey and Panel Data." *Journal of Money, Credit and Banking* 22 (2): 253–62.

Foote, C. L., K. Gerardi, L. Goette, and P. S. Willen. 2008. "Just the Facts: An Initial Analysis of Subprime's Role in the Housing Crisis." *Journal of Housing Economics* 17 (4): 291–305.

Gropp, R., J. K. Scholz, and M. J. White. 1997. "Personal Bankruptcy and Credit Supply and Demand." *Quarterly Journal of Economics* 112 (1): 217–51.

Hall, R. E., and F. S. Mishkin. 1982. "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households." *Econometrica* 50 (2): 461–81.

Hayashi, F. 1985. "The Effect of Liquidity Constraints on Consumption: A Cross-Sectional Analysis." *Quarterly Journal of Economics* 100 (1): 183–206.

Hubbard, R. G., and K. L. Judd. 1986. "Liquidity Constraints, Fiscal Policy, and Consumption." *Brookings Papers on Economic Activity* 17 (1, Spring): 1–60.

Jappelli, T. 1990. "Who Is Credit Constrained in the U.S. Economy?" *Quarterly Journal of Economics* 105 (1): 219–34.

Kennickell, A. B. 1991. "Imputation of the 1989 Survey of Consumer Finances: Stochastic Relation and Multiple Imputation." Working Paper, Board of Governors of the Federal Reserve System.

Kennickell, A. B., D. A. McManus, and R. L. Woodburn. 1996. "Weighting Design for the 1992 Survey of Consumer Finances." Working Paper, Board of Governors of the Federal Reserve System.

Keys, B. J., T. Mukherjee, A. Seru, and V. Vig. 2009. "Financial Regulation and Securitization: Evidence from Subprime Loans." *Journal of Monetary Economics* 56 (5): 700–20.

———. 2010. "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans." *Quarterly Journal of Economics* 125 (1): 307–62.

King, M. A., and L. Dicks-Mireaux. 1981. "Asset Holdings and the Life Cycle." NBER Working Paper No. 614.

———. 1982. "Asset Holdings and the Life Cycle." *Economic Journal* 92 (366): 247–67.

Lyons, A. C. 2003. "How Credit Access Has Changed Over Time for U.S. Households" *Journal of Consumer Affairs* 37 (2): 231–55.

Magri, S. 2007. "Italian Households' Debt: The Participation to the Debt Market and the Size of the Loan." *Empirical Economics* 33 (3): 401–26.

Manrique, J., and K. Ojah. 2004. “Credits and Non-Interest Rate Determinants of Loan Demand: A Spanish Case Study.” *Applied Economics* 36 (8): 781–91.

Mariger, R. P. 1987. “A Life-Cycle Consumption Model with Liquidity Constraints: Theory and Empirical Results.” *Econometrica* 55 (3): 533–57.

Mian, A., and A. Sufi. 2009a. “Household Leverage and the Recession of 2007 to 2009.” Paper presented at the 10th Jacques Polak Annual Research Conference, hosted by the International Monetary Fund, Washington D.C., November 5–6.

———. 2009b. “The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis.” *Quarterly Journal of Economics* 124 (4): 1449–96.

———. 2010. “House Prices, Home Equity-Based Borrowing, and the U.S. Household Leverage Crisis.” Chicago Booth Research Paper No. 09-20, Booth School of Business, University of Chicago.

Purnanandam, A. 2011. “Originate-to-Distribute Model and the Subprime Mortgage Crisis.” *Review of Financial Studies* 24 (6): 1881–1915.

Runkle, D. E. 1991. “Liquidity Constraints and the Permanent-Income Hypothesis: Evidence from Panel Data.” *Journal of Monetary Economics* 27 (1): 73–98.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

Zeldes, S. P. 1989. “Consumption and Liquidity Constraints: An Empirical Investigation.” *Journal of Political Economy* 97 (2): 305–46.