Volume 13, Number 4

December 2017

IJCB
INTERNATIONAL JOURNAL OF CENTRAL BANKING
Volume 13, Number 4
December 2017

IJCB

INTERNATIONAL JOURNAL OF CENTRAL BANKING

International Journal of Central Banking

Volume 13, Number 4 December 2017

*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**. Individual print subscriptions are available. Orders may be placed by phone (001 415 974 2035), via fax (001 415 974 2168), or by e-mail (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
Economic Research Department
Federal Reserve Bank of San Francisco
101 Market Street
San Francisco, CA 94105
USA

Phone: 001 415 974 2035
Fax: 001 415 974 2168
E-mail: editor@ijcb.org

ISSN: 1815-4654

A Global Trade Model for the Euro Area*

Antonello D'Agostino,[a] Michele Modugno,[b] and Chiara Osbat[c]
[a]Rokos Capital Management
[b]Federal Reserve Board
[c]European Central Bank

We propose a model for analyzing euro-area trade based on the interaction between macroeconomic and trade variables. First, we show that macroeconomic variables are necessary to generate accurate short-term trade forecasts; this result can be explained by the high correlation between trade and macroeconomic variables, with the latter being released in a more timely manner. Second, the model tracks well the dynamics of trade variables conditional on the path of macroeconomic variables during the Great Recession; this result makes our model a reliable tool for scenario analysis.

JEL Codes: F17, F47, C38.

1. Introduction

Understanding trade developments is a central issue for policy institutions as well as for the private sector since trade dynamics are important determinants of output growth and inflationary pressures coming from import prices. Having a model to infer current trade figures and future trade developments, conditional on macroeconomic scenarios, is important both for policy institutions, which form policy decisions, and for the private sector, which forms investment decisions.

*The opinions in this paper are those of the authors and do not necessarily reflect the views of the Board of Governors of the Federal Reserve System, the European Central Bank and the Eurosystem, and Rokos Capital Management. We thank anonymous referees for comments that much improved the clarity of the paper. Author e-mails: antonello.dagostino@rokoscapital.com, michele.modugno@frb.gov, chiara.osbat@ecb.int.

There are two main approaches to forecasting trade: large structural macro models and time-series models. Large structural models (e.g., Hervé et al. 2011; Riad et al. 2012) aim at understanding the economic mechanisms that generate trade dynamics, rather than at achieving the best possible forecasting performance. By contrast, time-series models (e.g., Jakaitiene and Dées 2012; Keck, Raubold, and Truppia 2009; Lin and Xia 2009; Yu, Wang, and Lai 2008) aim at building trade models with good forecasting properties.

Our work belongs in the time-series model literature, proposing a dynamic factor model that shows that exploiting the co-movement between macroeconomic variables and trade variables is essential for obtaining accurate short-term forecasts of trade variables. We use this model to infer future developments of trade variables given scenarios for macroeconomic variables and to quantify the effect on euro-area trade variables of changed macroeconomic conditions in euro-area trading partners.

In recent years, factor models have become a workhorse at central banks and international organizations for short-term forecasting of macroeconomic variables. The seminal paper of Giannone, Reichlin, and Small (2008) shows that factor models can handle easily a "ragged-edge" data structure, and that they produce very accurate short-term forecasts for U.S. real GDP. Several papers applied the same methodology for short-term forecasting of GDP, inflation, employment, and other variables for several countries; for a survey, see Bańbura et al. (2013) and Bańbura, Giannone, and Reichlin (2011). In this paper, we make use of a factor model estimated with the methodology proposed in Bańbura and Modugno (2014): they propose a maximum-likelihood estimation methodology based on a modification of the expectation maximization (EM) algorithm that allows to exploit data sets characterized by arbitrary patterns of missing data. Moreover, when using a maximum-likelihood estimation approach, it is straightforward to introduce restrictions on the parameters. This approach also allows to identify the nature of the unobserved factors.

We evaluate the model in a pseudo-out-of-sample simulation from January 2006 to April 2013: at each point in time we generate forecasts, we replicate the data availability as it was at that point in time, but we do not consider data revisions, given the scarce availability of real-time data.[1] We show that a factor model estimated on a panel of trade and macroeconomic data delivers accurate forecasts because it can fully exploit the co-movement in the panel and the earlier releases of the macroeconomic variables. The inclusion of real macroeconomic variables, confidence indicators, and prices improves the forecast accuracy over a model that exploits only trade information.

We also find, in contrast to Burgert and Dées (2009), but in line with Marcellino, Stock, and Watson (2003) for other euro-area macroeconomic variables, that the "bottom-up" forecast approach for euro-area exports and imports, which consists of producing forecasts for each set of partners separately and then aggregating them, delivers forecasts as good as those obtained with a "direct" approach that produces forecasts for the aggregate. This result is important, because it allows us to disentangle the contribution to the extra-euro-area forecast from different world regions, without sacrificing forecast performance.

We also run a natural experiment and generate the dynamics of trade variables during the Great Recession conditional on the realized path of macroeconomic variables. Results show that trade developments are well tracked: these results make our model a suitable tool for conditional scenario analyses.

The paper is organized as follows: section 2 describes the data and trade aggregation. Section 3 describes the model. Section 4 shows the forecasting results, while section 5 shows the conditional forecast exercise. Section 6 concludes.

2. Data

In this paper, we aim at forecasting monthly intra- and extra-euro-area import and export prices and volumes, vis-à-vis euro-area partners: Brazil, Russia, India, China, Japan, South Korea, Switzerland, Denmark, Sweden, the United Kingdom, Turkey, the United States, Canada, OPEC (Organization of the Petroleum Exporting Countries), and a residual called Rest of the World. There are

[1]For a discussion of the importance of the timing of data releases in nowcasting within the framework of a factor model, see Bańbura, Giannone, and Reichlin (2011).

in total sixty-eight trade data series (import and export volumes and prices from fourteen countries plus extra-euro-area, intra-euro-area, and the Rest of the World series). These data are produced by the statistics department of the European Central Bank (ECB).

In addition, for trade data, we exploit the predictive power of thirty macroeconomic variables selected on the basis of their availability: industrial production (IP) in manufacturing (from the euro area, Brazil, Canada, Denmark, the United Kingdom, India, Korea, Russia, Sweden, Turkey, and the United States); purchasing manager indexes (PMI) of new export orders (euro area, China, the United Kingdom, India, Japan, Korea, Russia, Turkey, and the United States); producer price indexes (PPI) in manufacturing (Brazil, Canada, Switzerland, Denmark, India, Japan, Korea, Russia, Sweden, and the United States), consumer price indexes (CPI) and PPI in the euro area; the real effective exchange rate of the euro area, deflated by CPI, vis-à-vis forty trading partners;[2] and the unemployment rate and retail sales in the euro area. The sample covers monthly observations from January 1995 to April 2013.[3] The data set is highly unbalanced due to the different publication lag of trade data, with various trading partners (some bilateral data are available before the aggregate), and macroeconomic data. There are also missing observations due to the initial date at which various macroeconomic variables become available.

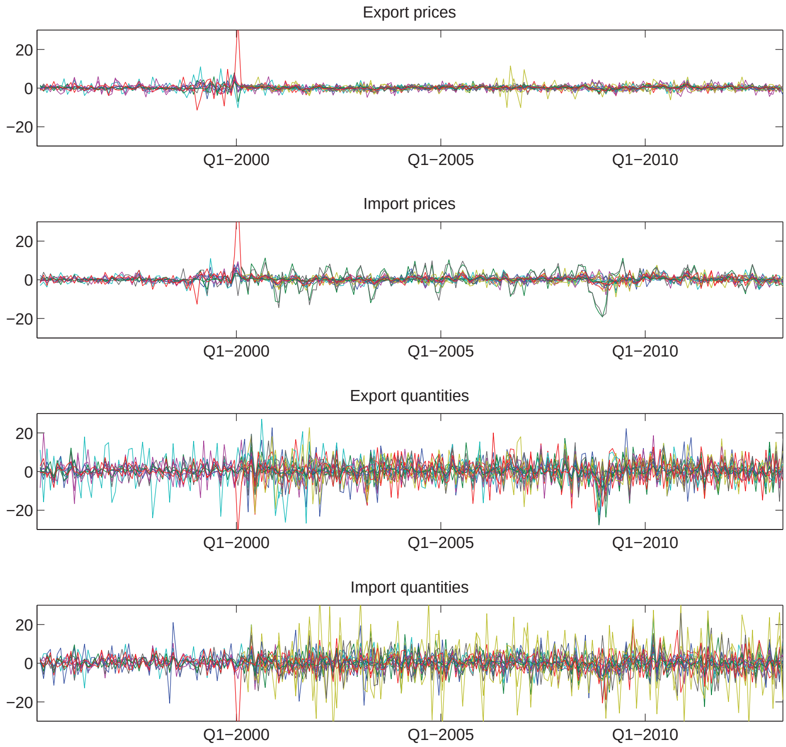
Figure 1 shows the monthly growth rates of the trade variables used in the paper: export prices, import prices, export volumes, and import volumes. The common feature of these data is the low persistence; in addition, volumes show higher volatility than prices.

Table 1 describes some baseline statistics: mean; standard deviation; absolute value of the autocorrelation coefficient, which is a measure of persistence; and the R^2, computed with a simple autoregressive model of order 1, which is a measure of predictability.

[2]Code EXR.M.Z65.EUR.ERC0.A at the ECB Statistical Data Warehouse, Exchange Rates.

[3]The data were downloaded on July 31, 2013.

Figure 1. Trade Variables



Note: Trade variables included in the analysis are month-on-month growth rates.

- Mean: On average, over the sample, month-on-month (MoM) extra- and intra-euro-area export price inflations both stand at 0.18, while extra- and intra-euro-area import price inflations are slightly higher at 0.29 and 0.21, respectively. Imports from Korea show negative price growth. Export volume growth, on average across the sample and across the fourteen countries, is 0.42. High growth is recorded in Russia, China, and Brazil at 0.96, 0.82, and 0.70, respectively. Extra-euro-area export growth is higher than intra-euro-area export growth, 0.36 versus 0.19. The average growth in export volumes across the fourteen countries is higher than that in import volumes (0.21).

Table 1. Descriptive Statistics of the Trade Data

	Export Prices				Import Prices				Export Volumes				Import Volumes			
	μ	σ	$\lvert\rho\rvert$	R^2	μ	σ	$\lvert\rho\rvert$	R^2	μ	σ	$\lvert\rho\rvert$	R^2	μ	σ	$\lvert\rho\rvert$	R^2
Brazil	0.03	1.19	0.20	0.04	0.33	2.25	0.22	0.05	0.70	8.56	0.28	0.08	0.14	5.71	0.19	0.04
Russia	0.17	1.11	0.25	0.06	0.64	5.33	0.37	0.14	0.96	6.55	0.35	0.12	0.20	7.21	0.50	0.25
India	0.10	1.45	0.34	0.12	0.13	1.75	0.16	0.03	0.59	7.76	0.31	0.10	0.67	5.55	0.37	0.14
China	0.28	2.23	0.27	0.07	0.08	2.47	0.04	0.00	0.82	7.87	0.40	0.16	1.01	5.17	0.35	0.12
Japan	0.11	1.98	0.16	0.03	0.10	1.39	0.05	0.00	0.17	5.85	0.50	0.25	−0.05	4.40	0.33	0.11
Korea	0.12	2.36	0.34	0.12	−0.04	1.98	0.02	0.00	0.56	7.86	0.41	0.17	0.28	12.72	0.52	0.27
Switzerland	0.19	1.18	0.22	0.05	0.25	1.99	0.34	0.11	0.22	4.40	0.48	0.23	0.17	5.31	0.51	0.26
Denmark	0.15	1.13	0.39	0.15	0.18	1.58	0.31	0.09	0.16	3.34	0.32	0.10	0.07	6.28	0.55	0.29
Sweden	0.14	1.40	0.29	0.08	0.15	1.21	0.21	0.05	0.28	4.68	0.40	0.16	0.20	3.90	0.39	0.15
United Kingdom	0.17	1.17	0.06	0.00	0.30	1.55	0.04	0.00	0.19	3.49	0.43	0.18	−0.07	3.30	0.44	0.19
Turkey	0.16	1.17	0.30	0.09	0.16	1.07	0.06	0.00	0.60	6.68	0.05	0.00	0.39	2.76	0.08	0.01
United States	0.22	2.01	0.27	0.08	0.25	2.08	0.07	0.00	0.30	5.00	0.43	0.18	0.08	3.20	0.27	0.07
Canada	0.12	2.01	0.34	0.11	0.19	2.58	0.13	0.02	0.31	5.94	0.43	0.18	−0.04	12.86	0.45	0.20
OPEC	0.16	1.05	0.21	0.05	0.62	5.15	0.43	0.18	0.69	6.77	0.54	0.29	−0.08	7.41	0.38	0.14
Extra	0.18	0.76	0.11	0.01	0.29	1.40	0.32	0.10	0.36	2.35	0.26	0.07	0.25	2.20	0.39	0.15
Intra	0.18	0.64	0.24	0.06	0.21	0.66	0.19	0.04	0.19	1.89	0.20	0.04	0.17	2.29	0.28	0.08
Rest of the World	0.28	2.94	0.02	0.00	0.46	3.74	0.15	0.02	0.23	5.46	0.38	0.15	0.10	5.50	0.30	0.09

Note: μ is the average, σ the standard deviation, $|\rho|$ (absolute value of) the autocorrelation coefficient of order one, and the R^2 statistic is computed with an autoregressive model of order one with the constant.

- Standard Deviation: Volumes display high volatility when compared with prices. Volatility is, on average, 5.35 and 5.43 for export and import volumes, respectively. Changes in price movements have instead more contained fluctuations; the average volatility is 1.46 and 2.17 for export and import prices inflation, respectively.
- Persistence: Persistence, measured as the absolute value of the first-order autocorrelation coefficient, is quite low for all of the series considered; on average, it is lower for prices (around 0.2) than for volumes (around 0.35).
- Predictability: The R^2 from an autoregressive model of order 1 can be interpreted as the (in-sample) percentage of the series variance that can be predicted. Trade series show a very low predictable component; the R^2, on average, is 0.06 and 0.05 for export and import prices, respectively, while it is slightly higher, 0.14 and 0.15, for export and import volumes. This makes the trade data hard to forecast, at least by using the traditional univariate time-series model. One possibility to improve the forecastability is then to use information embedded in the cross-sectional dimension. The idea put forward in the paper is to augment the panel of trade variables with a block of macro variables, which have some degree of forecastability and, at the same time, are cross-correlated with the trade block. This is explained in the methodology section.

2.1 Trade Data Aggregation

This paper proposes a model that can deliver accurate forecasts of volumes and prices of euro-area imports and exports. For the sake of simplicity, bilateral trade variables are grouped in the following geographical areas: Brazil, Russia, India, and China (BRIC); Japan and South Korea (Far East); Switzerland, Denmark, Sweden, the United Kingdom, and Turkey (Europe); and the United States and Canada (North America). The countries that are members of OPEC are grouped in the Rest of the World. Our data set includes aggregated intra- and aggregated extra-euro-area trade variables (import/export prices and volume). Such aggregated variables can be forecast directly or, alternatively, predicted by aggregating forecasts of each geographical area (the bottom-up approach). We show that the forecast accuracy of the bottom-up approach is superior or very close to the accuracy obtained by forecasting the aggregate series directly. This result is crucial, because it allows us to disentangle and measure the forecast contribution from geographical areas to the aggregate forecast (extra-euro area) without paying a price in terms of aggregate forecasting performance.

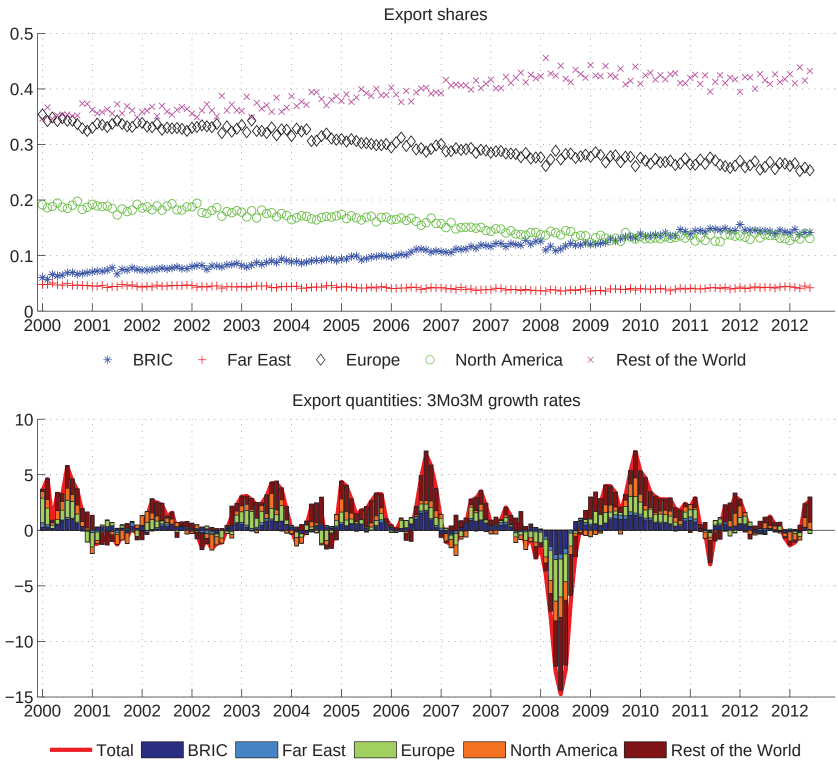
The aggregations are constructed using time-varying weights: weights for volumes are computed by rebasing volume indexes to the correspondent values as in 2000; weights for prices are computed using values in euro. Rebasing the volume indexes to values as in 2000 also allows us to construct volumes for the Rest of the World and their respective prices (recall that we measure prices by the implicit deflators obtained by dividing the values in euro by volume indexes).

Before moving to the forecasting exercise, we analyze the evolution of import and export shares over time; in addition, we show that the weighted average of the geographical growth rates accurately matches the growth rate of the aggregate series. The top panel in figure 2 shows the evolution of export shares for the five geographical areas. Different patterns can be observed over time: the export share to BRIC ($*$ symbol) increases constantly from 5 percent in 2000 to 15 percent at the end of the period. The export share to North America declines from values close to 20 percent in 2000 to around 14 percent in the beginning of 2013 ($\circ$ symbol). The same pattern can be observed for the export share to Europe: it declines from 35 percent in 2000 to just above 25 percent in 2013 ($\diamond$ symbol). The share of exports to the Rest of the World increases from 35 percent in 2000 to around 42 percent at the end of the sample period ($\times$ symbol). Finally, the export share of the Far East is relatively constant, just below 5 percent ($+$ symbol).

The bottom panel in figure 2 shows the three-month-on-three-month (3Mo3M) extra-euro-area exports growth rate (red straight line) and its breakdown by geographical areas (colored bars)[4]. Given the low persistence and high volatility of the data, we prefer including graphs and forecasts in 3Mo3M growth rates. This aggregation facilitates the interpretation of the results, because it removes the

[4]Colors in this figure and others in the paper can be seen in the online version available at www.ijcb.org.

Figure 2. Contribution to Extra-Euro-Area Export Volumes Growth by Geographical Breakdown

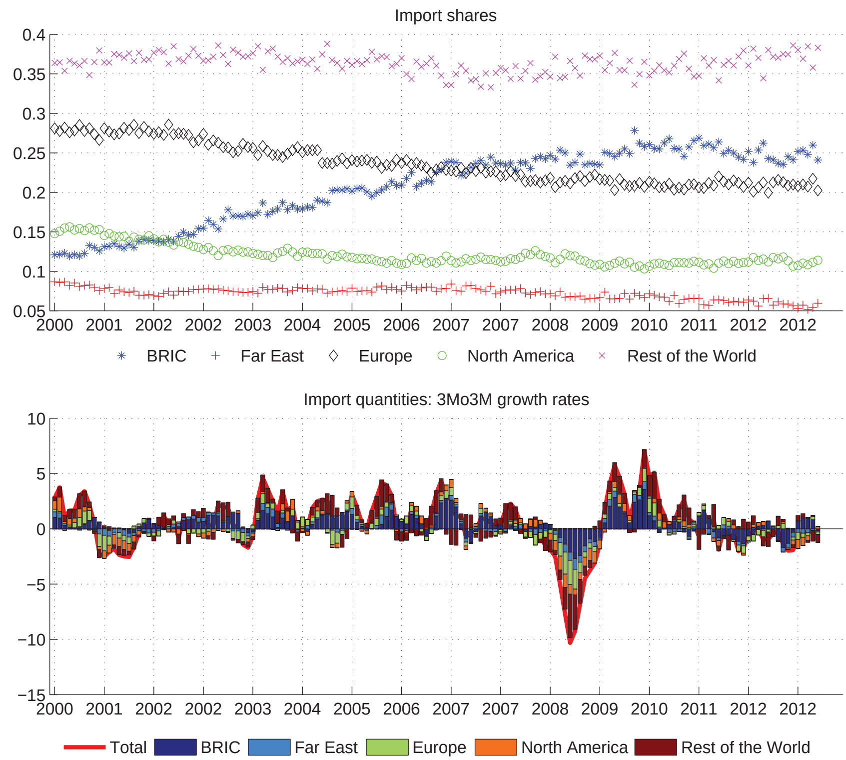


Notes: Top graph: extra-euro-area export shares by geographical areas. Bottom graph: three-month-on-three-month (3Mo3M) extra-euro-area export volume growth rate (red straight line) and the relative growth rate contributions by geographical areas (colored bars).

high-frequency noise component of the data. The forecasting model is estimated on month-on-month transformations.

First and most importantly, the aggregation by geographical growth rates reconstructs the aggregate extra-euro-area export series (blue line) quite well, which is important because we can retrospectively analyze the contribution of the single countries/areas to the aggregate growth rate. For example, in the Great Recession period, all the areas substantially contributed to the drop in trade, while in the last years of our sample the contribution of Europe is negligible.

Figure 3. Contribution to Extra-Euro-Area Import Volumes Growth by Geographical Breakdown



Notes: Top graph: extra-euro-area import shares by geographical areas. Bottom graph: three-month-on-three-month (3Mo3M) extra-euro-area import volume growth rate (red straight line) and the relative growth rate contributions by geographical areas (colored bars).

The top panel of figure 3 shows the evolution for the euro-area import shares from extra-euro-area countries. Euro-area imports from BRIC ($*$ symbol) grew from around 12 percent in 2000 to 25 percent in 2013. The euro-area import share from North America declined from 15 percent in 2000 to around 10 percent in 2013 ($\circ$ symbol). The euro-area import share from our Europe block dropped from around 28 percent in 2000 to 21 percent at the end of the sample period ($\diamond$ symbol). The euro-area import share from the Rest of the

World has remained fairly constant over time, around 36 percent (× symbol). The euro-area import share from our Far East block (+ symbol) has declined over time.

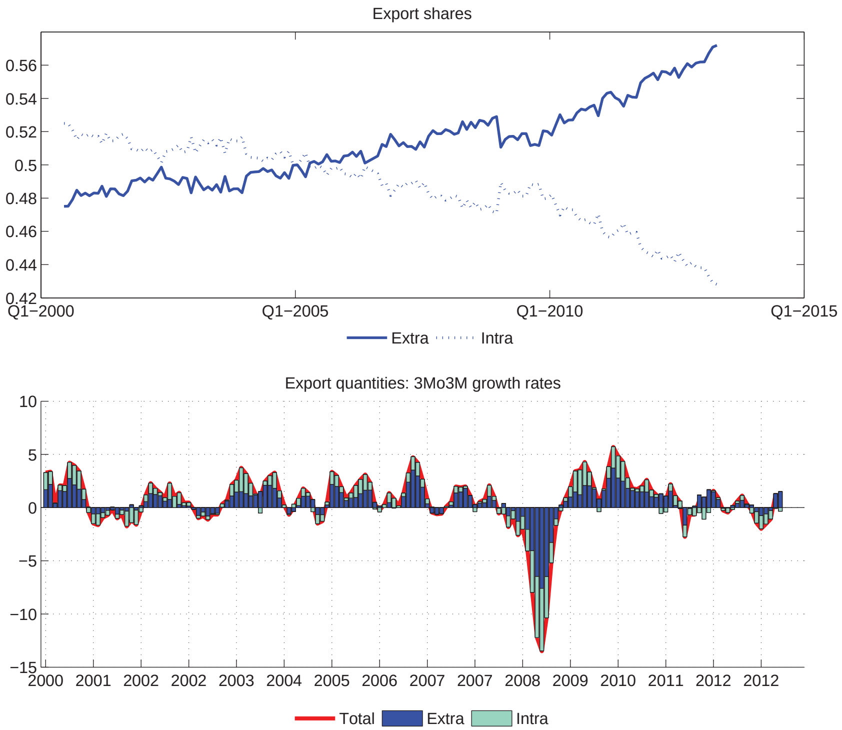
The bottom panel of figure 3 shows the 3Mo3M extra-euro-area import growth rate (blue straight line) and a geographical breakdown. The aggregation by countries/areas closely matches the aggregate extra-euro-area series, also in this case. During the Great Recession, imports from different areas declined. Since 2011, the import pattern has been less synchronized: positive growth contributions from some areas have been counterbalanced by negative contributions from other areas.

Figures 4 and 5 analyze these trade patterns for total euro-area exports and imports, dividing them into extra-euro area and intra-euro area. The share of extra-euro-area exports over total exports has increased over time from 48 percent in 2000 to around 57 percent in 2013 (figure 4). A less steep but similar trend is observed for extra-euro-area imports in figure 5. In 2005, for both imports and exports, the intra- and the extra-euro-area shares were equally split, at 50 percent each.

3. Econometric Framework

As shown in section 2, trade data are characterized by low persistence. In order to produce accurate forecasts of trade data, we exploit the cross-correlation among trade variables and in turn their correlation with other macroeconomic data, which are more timely than trade variables, i.e., they have a shorter publication delay. In order to exploit the cross-correlation between macroeconomic and trade variables, we use a dynamic factor model. Factor models can summarize the co-movement of a potentially large set of observable data with few common factors. As shown in Giannone, Reichlin, and Small (2008), factor models work well for forecasting a variable (GDP in their case) that is characterized by substantial publication lag but, at the same time, displays a strong correlation with other data characterized by shorter publication delay (surveys, industrial production, etc.). Given that trade variables are correlated with other macro variables but are published later than other macroeconomic

Figure 4. Contribution to Euro-Area Export Volumes Growth: Extra and Intra Decomposition



Notes: Top graph: extra- and intra-euro-area export shares. Bottom graph: three-month-on-three-month (3Mo3M) euro-area export quantity growth rate (red straight line) and the extra- and intra-euro-area relative growth rate contributions (colored bars).

variables, factor models are a natural tool to forecast them. The model is

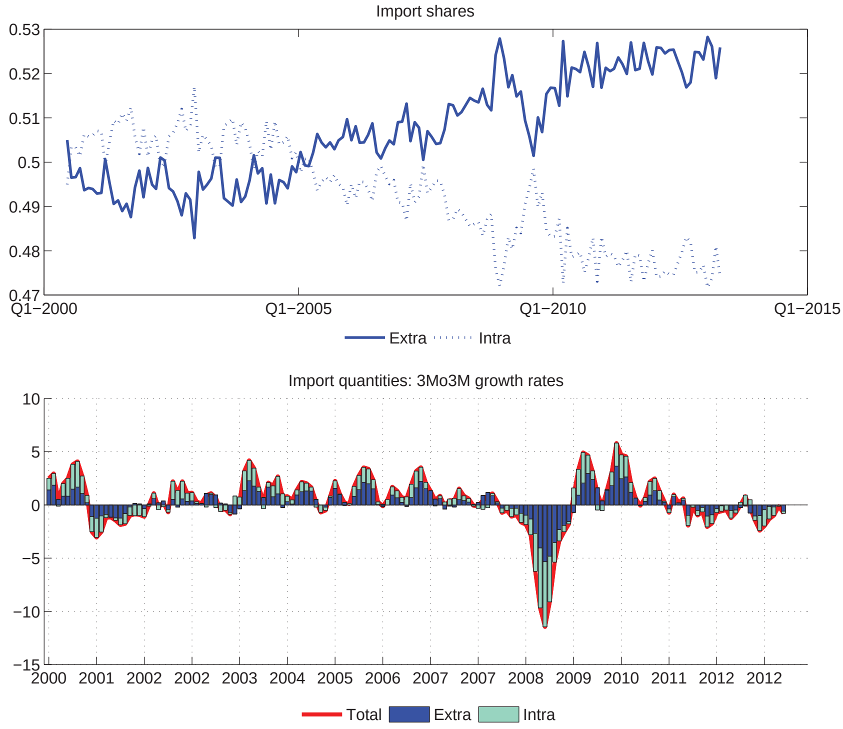
$$y_t = \Lambda f_t + \xi_t \tag{1}$$

$$f_t = A_1 f_{t-1} + \cdots + A_p f_{t-p} + u_t \qquad u_t \sim N(0, Q) \tag{2}$$

$$\xi_t = B\xi_{t-1} + \epsilon_t \qquad \epsilon_t \sim N(0, R), \tag{3}$$

where $y_t = [y_{1,t}, y_{2,t}, \ldots, y_{n,t}]'$, $t = 1, \ldots, T$ denote a stationary n-dimensional vector process with zero mean and unit variance. This

Figure 5. Contribution to Euro-Area Import Volumes Growth: Extra and Intra Decomposition



Notes: Top graph: extra- and intra-euro-area import shares. Bottom graph: three-month-on-three-month (3Mo3M) euro-area import volume growth rate (red straight line) and the extra- and intra-euro-area relative growth rate contributions (colored bars).

vector includes the observable data, i.e., trade and other macroeconomic variables; y_t depends on f_t, an r-dimensional vector of few unobserved common factors ($r << n$) and n idiosyncratic components $\xi_t = [\xi_{1,t}, \xi_{2,t}, \ldots, \xi_{n,t}]'$, which are uncorrelated with f_t at all leads and lags. Λ is an $n \times r$ matrix of factor loadings. It is also assumed that the common factors f_t follow a stationary vector autoregressive process of order p, where $A_1, \ldots, A_p$ are $r \times r$ matrices of lagged coefficients. We model the dynamics of the idiosyncratic components ξ_t as independent first-order autoregressive processes, i.e., the matrix B is diagonal; ϵ_t is normally distributed and cross-sectionally uncorrelated (the variance-covariance matrix R is diagonal), i.e., y_t follows an exact factor model.

3.1 Estimation

As described above, our data set has missing observations, not only at the end of the sample period (due to the different publication delays) but also at the beginning (due to the different time spans covered by the different series). In order to estimate the parameters of the model described by equations (1)–(3), given that we want to include restrictions on these parameters, and given the missing observations, we implement a maximum-likelihood algorithm. More precisely, we make use of the algorithm proposed by Bańbura and Modugno (2014), i.e., a modification of the expectation-maximization (EM) algorithm that allows to estimate the parameters of the model described by equations (1)–(3) with arbitrary patterns of missing observations.

The EM algorithm is a natural choice for dealing with the issues that arise when estimating parameters of a dynamic factor model. The first issue is that f_t, the vector of common factors, is unobserved, which implies that the maximum-likelihood estimates of the parameters are in general not available in closed form. Dempster, Laird, and Rubin (1977) introduced the EM algorithm as a general solution to problems for which latent states yield a likelihood function that is intractable. They propose to express the likelihood in terms of both observed and unobserved variables and iterating between two operations: (i) computing the expectation of the log-likelihood (sufficient statistics) conditional on the data using the parameter estimates from the previous iteration, and (ii) reestimating the parameters through the maximization of the expected log-likelihood. In the case of our model, this algorithm simplifies to an iteration between the two steps until convergence is achieved, while correcting at each step for the uncertainty associated with the estimation of the common factors (Shumway and Stoffer 1982; Watson and Engle 1983).

The second issue is due to the large number of series included in the panel. When n is large, the assumption of an exact factor structure, i.e., the matrix R is diagonal, can be the source of misspecification, given that some local cross-correlation can still survive after controlling for the common factors. However, Doz, Giannone, and Reichlin (2011) show that the effect of this misspecification on the estimation of the common factors is negligible when the sample size (T) and the cross-section (n) are large. They show that the factors extracted under the assumption of zero cross-correlation among the idiosyncratic components span the space of true factors from the model with weak cross-correlation among the idiosyncratic components. Moreover, they show that the estimator is feasible when n is large and easily implementable using the Kalman smoother and the EM algorithm as in traditional factor analysis.

3.2 Restrictions on the Parameters

One of the advantages of the maximum-likelihood approach, with respect to non-parametric methods based on principal components, is that it allows us to impose restrictions on the parameters in a relatively straightforward manner.

Bork (2009) and Bork, Dewachter, and Houssa (2009) show how to modify the maximization step of EM algorithm described by Watson and Engle (1983) in order to impose restrictions of the form $H_\Lambda \text{vec}(\Lambda) = \kappa_\Lambda$ for the model described in equations (1)–(3). Bańbura and Modugno (2014) show how those restrictions can be imposed in the presence of an arbitrary pattern of missing data.

We impose restrictions on the factor loadings matrix (Λ) in order to identify the factors in our model. The factor loadings are restricted to be equal to zero if the corresponding data series are not included in the group that identifies a factor. We assume there are four factors related to import prices, export prices, import volume, and export volume dynamics:

(i) f^1 is the factor capturing the co-movement among export volumes, euro-area trade partners' PMIs and industrial productions, and the real effective exchange rate.

(ii) f^2 is the factor capturing the co-movement among the real effective exchange rate, all import volumes, and the real euro-area macroeconomic variables, i.e., industrial production, retail sales, and the unemployment rate.

(iii) f^3 is the factor capturing the co-movement among the real effective exchange rate, all export prices, and the nominal euro-area macroeconomic variables, i.e., CPI and PPI.

(iv) f^4 is the factor capturing the co-movement among the real effective exchange rate, all import prices, and the euro-area trade partners' PPIs.

The exchange rate is the only variable that is included in the four factors. In table 4 of the appendix, we report the block structure.

In order to understand if the information content of macroeconomic variables increases the forecasting accuracy for trade variables, we compare the model described so far with a similar model that includes only trade variables. This second model is also characterized by four factors, defined as follows:

(i) f^1_{ot} is the factor capturing the co-movement among all export volumes.

(ii) f^2_{ot} is the factor capturing the co-movement among all import volumes.

(iii) f^3_{ot} is the factor capturing the co-movement among all export prices.

(iv) f^4_{ot} is the factor capturing the co-movement among all import prices.

The block structure for the trade variables is described in the appendix (table 5). Note that we do not impose any restriction to the transition equation (2): the factors interact with each other through the vector autoregressive process.

3.3 Forecasting

Forecasts are generated from the parameter estimates of the model described in equations (1)–(3) $\hat{\theta} = [\hat{\Lambda}, \hat{A}_1, \ldots, \hat{A}_p, \hat{B}, \hat{R}, \hat{Q}]$ and the data set Ω_v. The forecasts are defined as conditional expectations of the target variable $y_{i,t}$, obtained at time v, given the information set Ω_v. Notice that v refers to the point in time at which we produce the forecast and v can refer to any time frequency; we will assume that v is monthly, i.e., the forecast performance is evaluated every month. If $t < v$, we are backcasting; if $t = v$, we are nowcasting; and if $t > v$, we are forecasting. Forecasts are computed as

$$\mathbb{E}_{\hat{\theta}}\left[y_{i,t}|\Omega_v\right] = \hat{\Lambda}_{i\cdot}\mathbb{E}_{\hat{\theta}}\left[f_t|\Omega_v\right] + \mathbb{E}_{\hat{\theta}}\left[\xi_{i,t}|\Omega_v\right], \qquad y_{i,t} \notin \Omega_v, \quad (4)$$

where $\hat{\Lambda}_{i\cdot}$ denotes the i^{th} row of $\hat{\Lambda}$, the maximum-likelihood estimate of Λ. $\mathbb{E}_{\hat{\theta}}\left[f_t|\Omega_v\right]$ and $\mathbb{E}_{\hat{\theta}}\left[\xi_{i,t}|\Omega_v\right]$ are obtained by using Kalman filtering techniques, which can be applied to the state-space representation, described by equations (1)–(3), in order to derive backcasts, nowcasts, and forecasts of $y_{i,t}$. Finally, $\mathbb{E}_{\hat{\theta}}\left[\xi_{i,t}|\Omega_v\right] \neq 0$, given that, in our case, the idiosyncratic components follow a first-order autoregressive process.

4. Forecast Evaluation

We evaluate the forecast performance of our model via a pseudo-real-time out-of-sample simulation on the sample January 2006 to April 2013: the forecast horizon varies from -2 to 0, where -2 and -1 are the previous two months' and one month's backcasts, respectively, and 0 is the current month's nowcast. We produce these estimates every month, with a data set characterized by a "ragged-edge" structure that mimics the information available at the end of each month. The vintage of data on which our estimates are based was downloaded on July 31, 2013. For example, let us assume that we start the forecast evaluation on January 31, 2006. On this day, trade data relative to Denmark, Sweden, and Great Britain are available until October 2005, while all the other trade data are available up to November 2005. Macroeconomic variables are more timely: PMIs and IPs of Russia and the United States are available for December 2005, while the IPs for all the other non-euro-area countries are available up to November 2005. PPIs are all available up to December 2005, except for Brazil, for which PPI data are available up to November 2005. Euro-area macroeconomic data are all available up to November 2005, but the exchange rate is available up to December 2005. This structure is exactly replicated at each forecast iteration to implement a pseudo-real-time exercise.

Forecasts are evaluated by the mean squared forecast error (MSFE) statistic defined as

$$MSFE_{t_0}^{t_1} = \frac{1}{t_1 - t_0 + 1} \sum_{t=t_0}^{t_1} \left(\hat{Y}_{t+h|t} - Y_{t+h} \right)^2, \tag{5}$$

where $\hat{Y}_{t+h|t}$ is the backcast or nowcast ($h = -2, h = -1$, and $h = 0$) of the target variables and Y_{t+h} are the ex post realized values; t_0 and t_1 are the starting and ending forecast evaluation periods.

The forecasting exercise has two aims: first, to understand the marginal forecasting power of the macro variables and, second, to compare the relative performance of the aggregate versus disaggregate forecast.

Results are reported as the ratio of the MSFE generated by the proposed models to the MSFE obtained with a benchmark naïve model, which is the constant growth model. A ratio smaller than 1 indicates that the factor model improved on the benchmark.

Table 2 compares the results for the two factor models without macro variables (panel A) and with macro variables (panel B). All the forecasts are computed by the bottom-up approach, that is, by aggregating forecasts from the different geographical areas. The factor model with only trade variables displays a much better two-month backcast performance ($h = -2$) compared with the naïve model (the relative MSFE is smaller than 1). The performance is in line with that of the benchmark for the one-month backcast ($h = -1$), is slightly worse for export and import prices (1.07 and 1.08, respectively), and is slightly better for import and export volume (0.94 and 0.98, respectively). The performance deteriorates for the import and export prices nowcast ($h = 0$) and is identical to that of the benchmark model for export volumes; it improves for import volumes by 10 percent with respect to the naïve model.

When we include macro variables in our model, there is a generalized improvement of the forecasting performance with respect to both the naïve benchmark and the model with only trade variables. Relative MSFEs in panel B are always smaller than those in panel A, the only exceptions being export and import volumes for $h = 0$, which are slightly higher.

Panel C in table 2 compares the relative performance of disaggregated forecasts with the aggregated forecasts. The latter is generated

Table 2. Relative MSFE: Extra-Euro-Area Trade (using block structure)

	Export Prices	Import Prices	Export Volumes	Import Volumes
A. Trade Variables Only				
$h = -2$	0.23	0.36	0.51	0.40
$h = -1$	1.07	1.08	0.98	0.94
$h = 0$	1.02	1.16	1.00	0.90
B. Trade and Macro Variables				
$h = -2$	0.27	0.16	0.43	0.37
$h = -1$	0.86	0.68	0.94	0.97
$h = 0$	0.81	0.87	1.16	0.99
C. Disaggregate vs. Aggregate Forecasts				
$h = -2$	0.82	1.06	0.87	1.21
$h = -1$	1.01	1.18	0.94	0.99
$h = 0$	1.02	1.06	0.99	1.01

Notes: Panel A reports the relative mean square forecast error (MSFE) between the MSFE obtained with a baseline constant growth model (denominator) and that obtained with the factor model computed on a panel with only trade variables (numerator). In panel B the factor model includes trade and macro variables; the numbers report the relative MSFE as in panel A. Panel C shows the ratio of the MSFE obtained aggregating the forecast for each single area to the MSFE obtained forecasting directly the aggregated series. A ratio below 1 indicates that the model at the numerator has a more accurate forecasting performance. $h = -2, h = -1$, and $h = 0$ denote the forecast horizons; in this case they refer to two-months-ago backcast, one-month-ago backcast, and nowcast, respectively. The evaluation period is January 2006 to April 2013.

by computing the predictions of extra-euro-area trade directly using the aggregate series. The results show that the performance of the bottom-up approach is similar to that of the direct forecast. On average, the relative MSFE statistics are around 1. This result is important, because it allows us to decompose the forecast contributions to the aggregate series by geographical areas.

The contribution of the macro variables is stronger in forecasting trade prices than in forecasting quantities. This can be explained by looking at the composition of the set of macro variables: theory would predict that foreign producer prices and exchange rates have an immediate and direct impact on import prices, and likewise for the impact of euro-area prices on export prices. The impact of foreign industrial production and growth on the export quantities demanded from the euro area would be mediated by relative prices and characteristics of foreign demand and it would manifest itself with a longer lag.[5]

What is the impact of imposing the block structure? Table 3 shows the forecast performance of the model when the block structure is removed. When the block structure is removed, the model does worse on import volumes both with and without the use of macro variables, except in one case. The striking improvement due to the block structure comes in forecasting prices, especially export prices, in the case of the disaggregate forecast. This can be expected in light of the intuition above: trade prices are affected by costs and markups, which in the data set are represented by producer price indexes and exchange rates. Unless this basic economic intuition is imposed by the block structure, the signal on prices is made more noisy and the performance decays substantially.

Figures 6–9 show the forecast results. Panels 1–3 in figure 6 show the backcast performance (panels 1 and 2 for $h = -2$ and $h = -1$, respectively) and the nowcast performance (panel 3) for 3Mo3M exports prices. The colored bars refer to the contribution of a geographical area to the aggregate extra-euro-area export prices series (blue dashed line). The red line shows the 3Mo3M realized value. The export price forecasts are quite accurate and track the realized series relatively well; the performance tends to deteriorate over the forecast horizon, from $h = -2$ to $h = 0$. The relative forecast contributions show that the forecast deflation in imports prices, over the last part of the sample, is mainly due to the North America forecast.

Figure 7 is similar to figure 6: it shows the forecast patterns for extra-euro-area export volumes. In this case, the forecasts are rather accurate, that is, they capture the great trade collapse in the middle of the sample, and, as expected, the accuracy deteriorates with the

[5] See, among others, Cooke (2014) and Lewis (2013).

Table 3. Relative MSFE: Extra-Euro-Area Trade (without block structure)

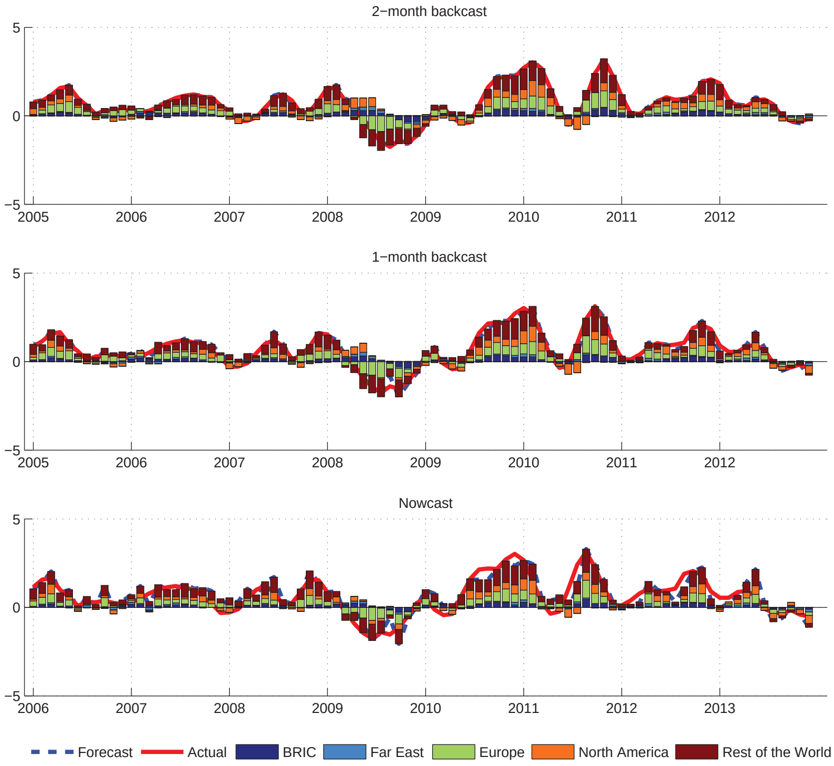
	Export Prices	Import Prices	Export Volumes	Import Volumes
A. Trade Variables Only				
$h = -2$	0.35	0.39	0.59	0.48
$h = -1$	1.05	1.03	1.04	0.98
$h = 0$	0.99	1.08	0.99	0.92
B. Trade and Macro Variables				
$h = -2$	0.33	0.31	0.55	0.44
$h = -1$	0.90	0.71	0.82	0.87
$h = 0$	0.69	0.82	1.04	1.02
C. Disaggregate vs. Aggregate Forecasts				
$h = -2$	4.96	1.76	0.84	1.25
$h = -1$	1.80	1.33	1.02	0.07
$h = 0$	1.37	1.03	1.05	1.04

Notes: Panel A reports the relative mean square forecast error (MSFE) between the MSFE obtained with a baseline constant growth model (denominator) and that obtained with the factor model computed on a panel with only trade variables (numerator). In panel B the factor model includes trade and macro variables; the numbers report the relative MSFE as in panel A. Panel C shows the ratio of the MSFE obtained aggregating the forecast for each single area to the MSFE obtained forecasting directly the aggregated series. A ratio below 1 indicates that the model at the numerator has a more accurate forecasting performance. $h = -2, h = -1$, and $h = 0$ denote the forecast horizons; in this case they refer to two-months-ago backcast, one-month-ago backcast, and nowcast, respectively. The evaluation period is January 2006 to April 2013.

forecast horizon. In terms of forecast contribution, the Rest of the World and Europe are the main components of the aggregate series; forecast contributions from different geographical areas tend to comove closely. However, this pattern is broken over the last part of the sample, where contributions from different areas to the aggregate forecast seem to be more erratic.

Figure 8 shows the results for import prices. Forecasts track the realized series well. Contributions to the aggregate forecasts tend to

Figure 6. Export Prices

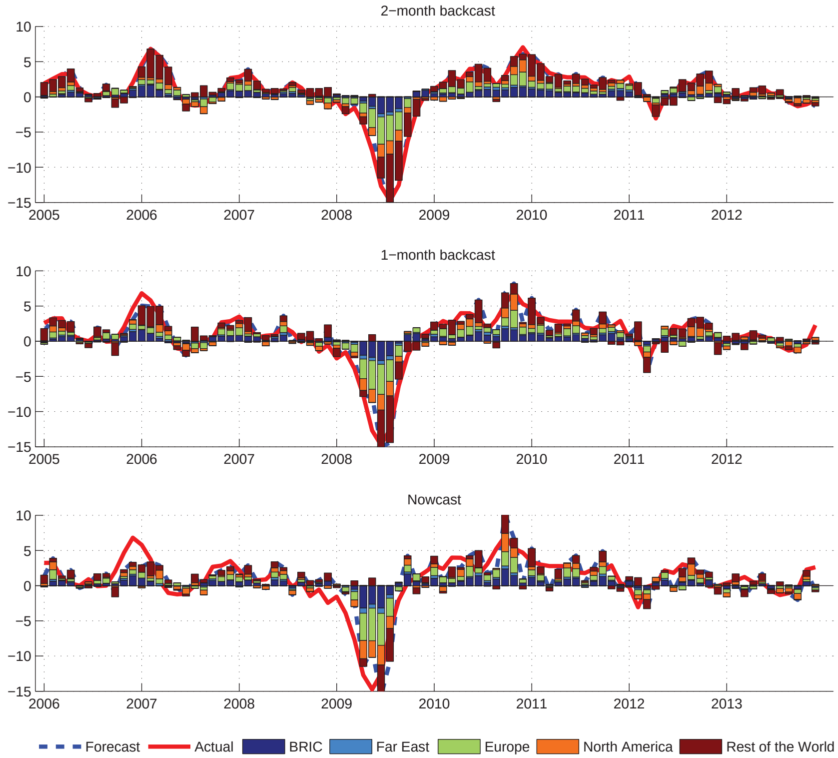


Notes: Three-month-on-three-month (3Mo3M) growth rates—extra-euro-area export prices forecasts. The dashed blue line is the aggregated extra-euro-area forecast. The colored bars are the forecast contributions from different geographical areas. The red straight line is the ex post realized value. Top graph: two-month backcast; middle graph: one-month backcast; bottom graph: nowcast.

co-move. The most important components for the aggregate dynamics are import price developments from BRIC and the Rest of the World.

Finally, figure 9 displays the results for import volumes. The overall performance looks quite good. The main contributors to the overall figure are, as in the case of import prices, BRIC and the Rest of the World. Imports from Europe played a non-trivial role in the Great Trade Collapse.

Figure 7. Export Volumes

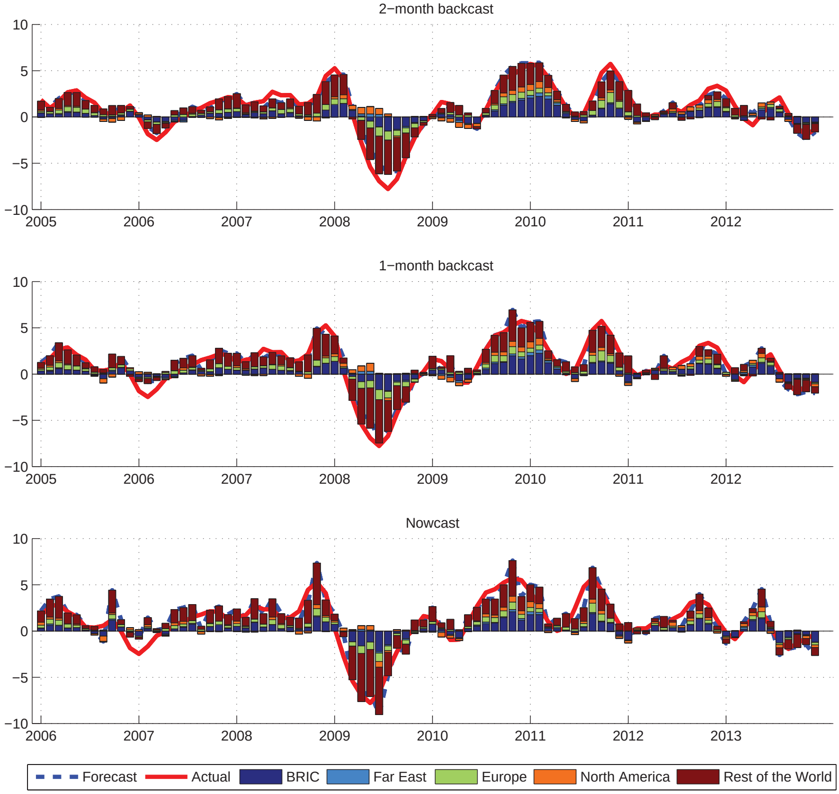


Notes: Three-month-on-three-month (3Mo3M) growth rates—extra-euro-area volumes forecasts. The dashed blue line is the aggregated extra-euro-area forecast. The colored bars are the forecast contributions from different geographical areas. The red straight line is the ex post realized value. Top graph: two-month backcast; middle graph: one-month backcast; bottom graph: nowcast.

5. Conditional Forecast

The econometric model specified in this paper can also be used to produce conditional forecasts, i.e., to evaluate the dynamics of a target variable conditional on the future path of some other variables. Conditional forecasting exercises have become more and more common in the macroeconometric literature and in policy circles to assess the stability of standard economic relationships (see, e.g., Dotsey, Fujita, and Stark 2011 and Giannone et al. 2014 for use

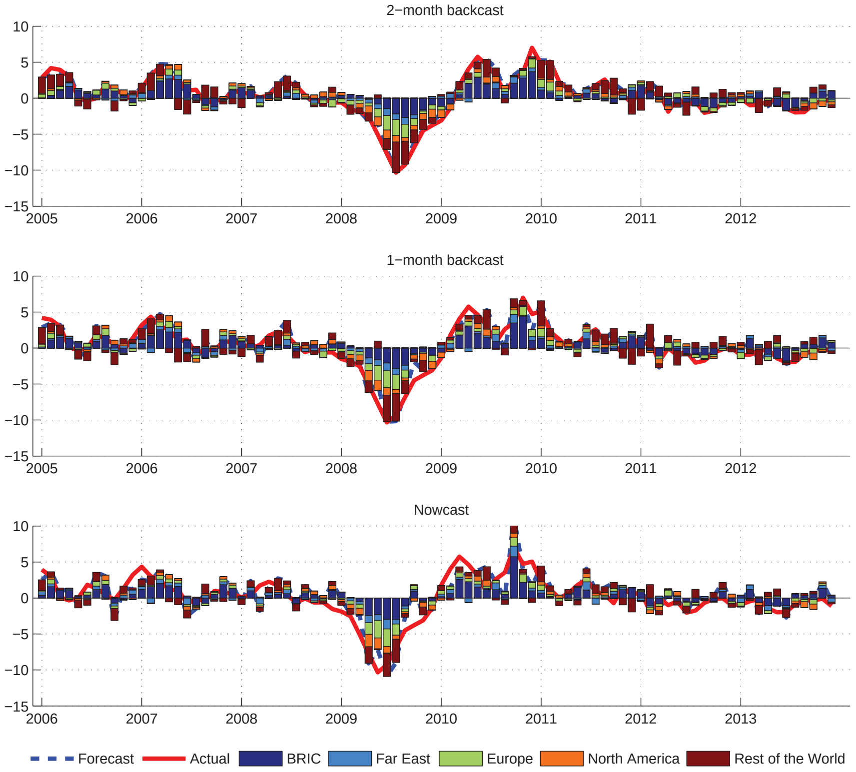
Figure 8. Import Prices



Notes: Three-month-on-three-month (3Mo3M) growth rates—extra-euro-area import prices forecasts. The dashed blue line is the aggregated extra-euro-area forecast. The colored bars are the forecast contributions from different geographical areas. The red straight line is the ex post realized value. Top graph: two-month backcast; middle graph: one-month backcast; bottom graph: nowcast.

of conditional forecasts to assess the stability of the Phillips curve) or to assess risks underlying forecasts, as in Baumeister and Kilian (2014). The state-space formulation of the factor model provides a natural framework to address this kind of exercise (see Bańbura, Giannone, and Lenza 2015). We consider a conditional forecast to examine how reliable our model is for producing trade data paths conditional on macroeconomic variables.

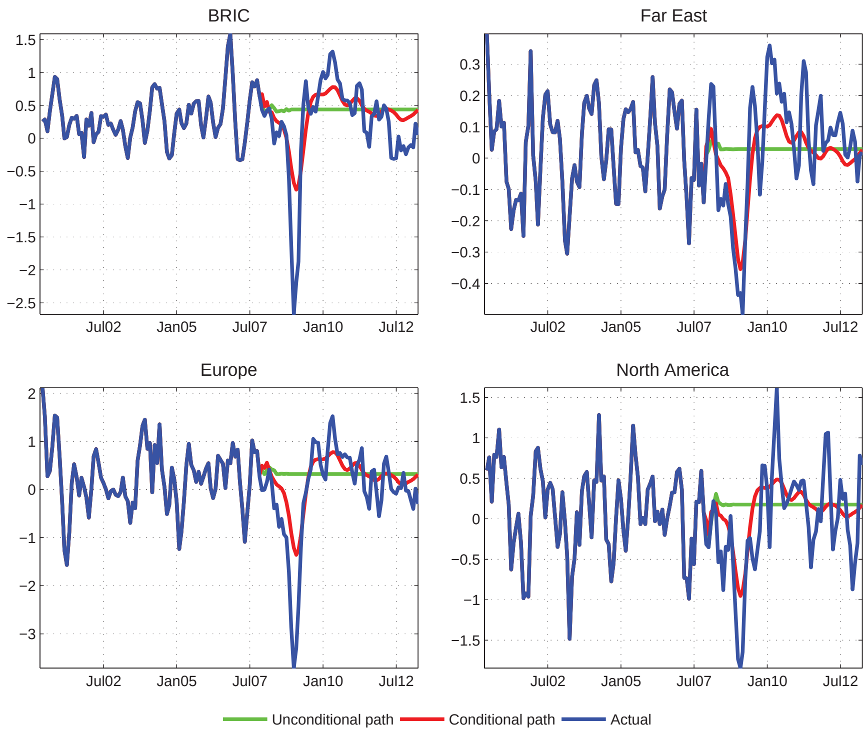
Figure 9. Import Volumes



Notes: Three-month-on-three-month (3Mo3M) growth rates—extra-euro-area import volumes forecasts. The dashed blue line is the aggregated extra-euro-area forecast. The colored bars are the forecast contributions from different geographical areas. The red straight line is the ex post realized value. Top graph: two-month backcast; middle graph: one-month backcast; bottom graph: nowcast.

In order to assess the reliability of our model to generate paths of trade variables conditional on macro variables, we conduct a natural experiment. Namely, we estimate the parameters with data available until December 2007 and then "feed" the Kalman filter with those parameters and with the observed macro variables from January 2008 to April 2013 to generate the conditional path of trade variables.

Figure 10. Conditional Path for Export Volumes

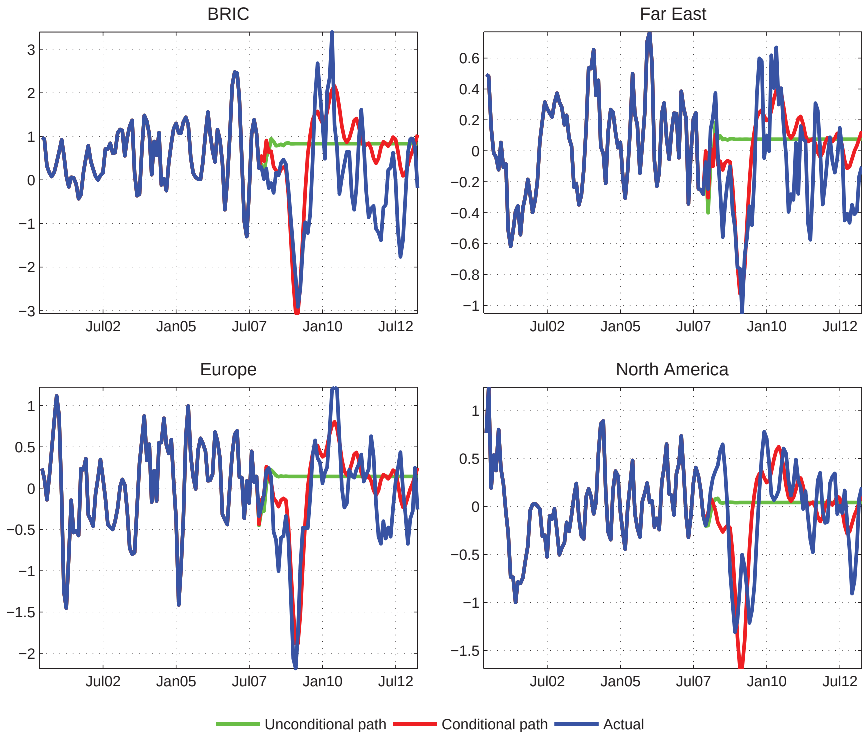


Notes: Simulated extra-euro-area export volumes path over the sample November 2007 to April 2013; trade dynamics are conditional on the realized macro values from November 2007 to April 2013 and on the parameters estimated with the data available at the end of December 2007. The four subplots refer to different geographical areas. Values are in three-month-on-three-month (3Mo3M) growth rates.

This exercise is informative: by comparing the conditional forecasts of the trade variables with their realized values, we provide an indirect measure of the importance of macro variables in driving trade dynamics. In addition, the exercise also provides an indication of the ability of the model to assess the effect of macroeconomic scenarios on the trade variables.

Figures 10 and 11 show the results for export and import volumes, respectively, and for the four different geographical areas (BRIC, Far East, Europe, and North America). Results for the Rest

Figure 11. Conditional Path for Import Volumes

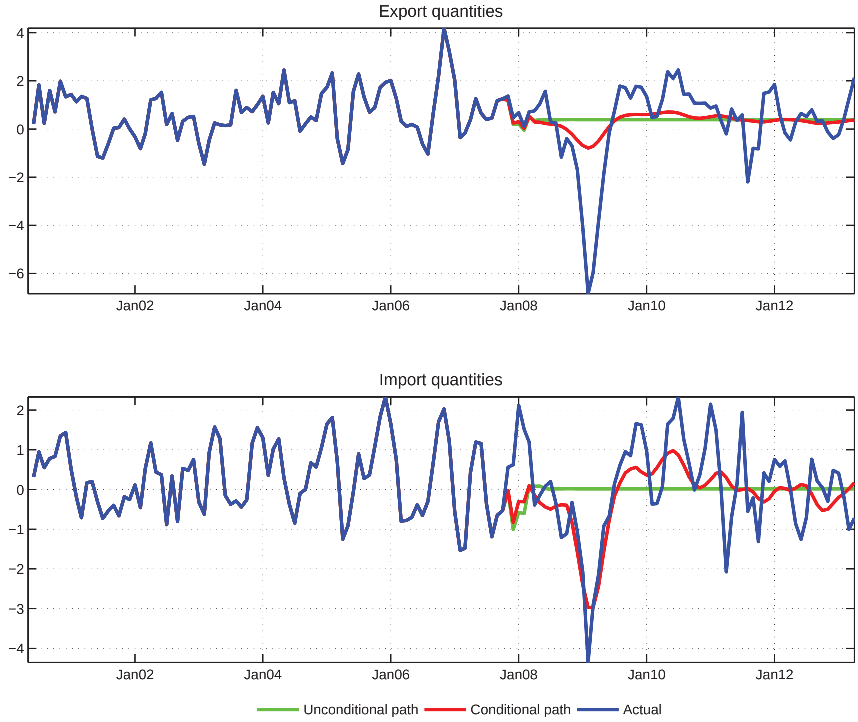


Notes: Simulated extra-euro-area import volumes path over the sample November 2007 to April 2013; trade dynamics are conditional on the realized macro values from November 2007 to April 2013 and on the parameters estimated with the data available at the end of December 2007. The four subplots refer to different geographical areas. Values are in three-month-on-three-month (3Mo3M) growth rates.

of the World block are reported in figure 12.[6] In all the charts, the red line refers to the 3Mo3M conditional growth rate dynamics, the blue line is the 3Mo3M ex post realized values, and the green line is the 3Mo3M unconditional path. The first panel of figure 10 shows the conditional path of (3Mo3M) exports volumes to BRIC (red line). In general, the simulated series (red line) track the actual

[6]We keep the trade weights constant at December 2007 values to aggregate the series from January 2008 to April 2013.

Figure 12. Conditional Path for Export and Import Volumes—Rest of the World



Notes: Simulated extra-euro-area export and import volumes paths over the sample November 2007 to April 2013; trade dynamics are conditional on the realized macro values from November 2007 to April 2013 and on the parameters estimated with the data available at the end of December 2007. The four subplots refer to different geographical areas. Values are in three-month-on-three-month (3Mo3M) growth rates.

path (blue line) well; they do not fully capture the depth of the drop in exports during 2009, but nonetheless show a substantial decline. Export dynamics to the Far East (top-right panel) are also well captured by the conditional series (red line); both the drop in 2009 and the general contour are well fitted. The two bottom panels show the conditional paths of exports to our Europe block and North America; the two series describe trade dynamics well, even if the trade collapse in 2009 is not fully captured.

Figure 11 shows the conditional paths of import volumes. The simulation exercise shows that the import drop in 2009 is well fitted in the four geographical areas; the general pattern is also well tracked, although there is a slight upward bias for the BRIC series (top-left panel) and the Far East series (top-right panel).

Figure 12 shows the results for the Rest of the World. The top panel displays the results for exports. The conditional series (red line) exhibits a poor fit; it does not co-move much with the observed values (blue line). This result can be explained by the lack of macroeconomic series in our data set for this "residual" region: macroeconomic data are a good proxy of the external demand component, which is correlated to exports.

The bottom panel in figure 12 shows the simulation for imports (red line). In this case, the conditional forecast displays a quite accurate fit. It captures the drop in 2009 and the general contour well.

6. Conclusions

In this paper, we study and forecast trade dynamics in the euro area. We use the factor model proposed by Bańbura and Modugno (2014). This model is a flexible tool to extract information from data sets characterized by arbitrary patterns of missing data. Furthermore, it allows for restrictions on the parameter space, which is essential for the identification of the factors in the model.

We focus on backcasting and nowcasting extra-euro-area import volumes, import prices, export price, export volumes, and their geographical subcomponents. In a pseudo-out-of-sample evaluation exercise starting in January 2006, we show that the model with trade and macro variables improves on the forecastability of a model with only trade variables. The more timely information of macro variables delivers an improvement in forecast accuracy.

In addition, we show that aggregating forecasts from euro-area trading regions (bottom-up approach) delivers predictions as accurate as those obtained by forecasting directly the extra-euro-area series ("direct" approach). This result is important, because it allows us to disentangle the contribution to the extra-euro-area trade forecast of different world regions.

Finally, we set up a counterfactual exercise where we show that macro variables track trade dynamics well; this result implies that future trade paths can be inferred by future macro paths, which are more predictable.

Appendix. Extra Tables

Table 4. Loading Restrictions on Macro Variables

	f^1	f^2	f^3	f^4
U.K. PMI	1	0	0	0
India PMI	1	0	0	0
Japan PMI	1	0	0	0
Korea PMI	1	0	0	0
Russia PMI	1	0	0	0
Turkey PMI	1	0	0	0
U.S. PMI	1	0	0	0
Brazil IP	1	0	0	0
Canada IP	1	0	0	0
U.K. IP	1	0	0	0
India IP	1	0	0	0
Japan IP	1	0	0	0
Korea IP	1	0	0	0
Russia IP	1	0	0	0
Sweden IP	1	0	0	0
Turkey IP	1	0	0	0
U.S. IP	1	0	0	0
Brazil PPI	0	0	0	1
Canada PPI	0	0	0	1
Denmark PPI	0	0	0	1
India PPI	0	0	0	1
Japan PPI	0	0	0	1
Sweden PPI	0	0	0	1
U.S. PPI	0	0	0	1
Euro-Area IP	0	1	0	0
Euro-Area RS	0	1	0	0
Euro-Area UR	0	1	0	0
Euro-Area REER	1	1	1	1
Euro-Area CPI	0	0	1	0
Euro-Area PPI	0	0	1	0

Notes: Factor loadings structure on macro variables. PMI stands for purchasing manager index, IP for industrial production, PPI for producer price index, RS for retail sales, UR for unemployment rate, REER for real effective exchange rate, and CPI for consumer price index. 1 indicates that there are no restrictions, 0 that the loading is restricted to zero.

Table 5. Loading Restrictions for the Model with only Trade Variables

	Export Prices				Import Prices				Export Volumes				Import Volumes			
	f^1	f^2	f^3	f^4	f^1	f^2	f^3	f^4	f^1	f^2	f^3	f^4	f^1	f^2	f^3	f^4
Brazil	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
Russia	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
India	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
China	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
Japan	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
Korea	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
Switzerland	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
Denmark	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
Sweden	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
United Kingdom	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
Turkey	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
United States	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
Canada	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
OPEC	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
Extra	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
Intra	0	0	1	0	0	0	1	1	1	1	0	0	0	1	0	0
Rest of the World	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0

Note: Factor loadings structure; 1 indicates that there are no restrictions, 0 that the loading is restricted to zero.

References

Bańbura, M., D. Giannone, and M. Lenza. 2015. "Conditional Forecasts and Scenario Analysis with Vector Autoregressions for Large Cross-Sections." *International Journal of Forecasting* 31 (3): 739–56.

Bańbura, M., D. Giannone, M. Modugno, and L. Reichlin. 2013. "Now-casting and the Real-Time Data Flow." In *Handbook of Economic Forecasting,* Vol. 2, Part A, ed. G. Elliott and A. Timmermann, 195–236. Elsevier.

Bańbura, M., D. Giannone, and L. Reichlin. 2011. "Nowcasting." In *Oxford Handbook on Economic Forecasting*, ed. M. P. Clements and D. F. Hendry. Oxford University Press.

Bańbura, M., and M. Modugno. 2014. "Maximum Likelihood Estimation of Factor Models on Datasets with Arbitrary Pattern of Missing Data." *Journal of Applied Econometrics* 29 (1): 133–60.

Baumeister, C., and L. Kilian. 2014. "Real-Time Analysis of Oil Price Risks Using Forecast Scenarios." *IMF Economic Review* 62 (1): 119–45.

Bork, L. 2009. "Estimating US Monetary Policy Shocks Using a Factor-Augmented Vector Autoregression: An EM Algorithm Approach." CREATES Research Paper No. 2009-11, School of Economics and Management, University of Aarhus.

Bork, L., H. Dewachter, and R. Houssa. 2009. "Identification of Macroeconomic Factors in Large Panels." CREATES Research Paper No. 2009-43, School of Economics and Management, University of Aarhus.

Burgert, M., and S. Dées. 2009. "Forecasting World Trade: Direct Versus "Bottom-Up" Approaches." *Open Economies Review* 20 (3): 385–402.

Cooke, D. 2014. "Monetary Shocks, Exchange Rates, and the Extensive Margin of Exports." *Journal of International Money and Finance* 41 (March): 128–45.

Dempster, A., N. Laird, and D. Rubin. 1977. "Maximum Likelihood Estimation from Incomplete Data." *Journal of the Royal Statistical Society: Series B* 39 (1): 1–38.

Dotsey, M., S. Fujita, and T. Stark. 2011. "Do Phillips Curves Conditionally Help to Forecast Inflation?" Working Paper No. 11-40, Federal Reserve Bank of Philadelphia.

Doz, C., D. Giannone, and L. Reichlin. 2011. "A Two-Step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering." *Journal of Econometrics* 164 (1): 188–205.

Giannone, D., M. Lenza, D. Momferatou, and L. Onorante. 2014. "Short-Term Inflation Projections: A Bayesian Vector Autoregressive Approach." *International Journal of Forecasting* 30 (3): 635–44.

Giannone, D., L. Reichlin, and D. Small. 2008. "Nowcasting: The Real-Time Informational Content of Macroeconomic Data." *Journal of Monetary Economics* 55 (4): 665–76.

Hervé, K., N. Pain, P. Richardson, F. Sédillot, and P.-O. Beffy. 2011. "The OECD's New Global Model." *Economic Modelling* 28 (1–2): 589–601.

Jakaitiene, A., and S. Dées. 2012. "Forecasting the World Economy in the Short Term." *The World Economy* 35 (3): 331–50.

Keck, A., A. Raubold, and A. Truppia. 2009. "Forecasting International Trade: A Time Series Approach." *OECD Journal: Journal of Business Cycle Measurement and Analysis* 2009 (2): 157–76.

Lewis, L. 2013. "Menu Costs, Trade Flows, and Exchange Rate Volatility." 2013 Meeting Paper No. 313, Society for Economic Dynamics.

Lin, L., and M. Xia. 2009. "A Novel Fuzzy Evolutionary Algorithm for Forecasting of International Trade in Dynamical Environment." In *2009 IITA International Conference on Services Science, Management and Engineering*, 325–28. Zhangjiajie, China: IEEE Computer Society.

Marcellino, M., J. H. Stock, and M. W. Watson. 2003. "Macroeconomic Forecasting in the Euro Area: Country Specific versus Area-Wide Information." *European Economic Review* 47 (1): 1–18.

Riad, N., L. Errico, C. Henn, C. Saborowski, M. Saito, and J. Turunen. 2012. "Changing Patterns of Global Trade." IMF Departmental Paper No. 12/01 (January 9).

Shumway, R., and D. Stoffer. 1982. "An Approach to Time Series Smoothing and Forecasting using the EM Algorithm." *Journal of Time Series Analysis* 3 (4): 253–64.

Watson, M. W., and R. F. Engle. 1983. "Alternative Algorithms for the Estimation of Dynamic Factor, Mimic and Varying

Coefficient Regression Models." *Journal of Econometrics* 23 (3): 385–400.

Yu, L., S. Wang, and K. K. Lai. 2008. "Forecasting China's Foreign Trade Volume with a Kernel-Based Hybrid Econometric-AI Ensemble Learning Approach." *Journal of Systems Science and Complexity* 21 (1): 1–19.

Optimal Monetary Policy with Nominal Rigidities and Lumpy Investment*

Tommy Sveen[a] and Lutz Weinke[b]
[a]BI Norwegian Business School
[b]Humboldt-Universität zu Berlin

New Keynesian theory generally abstracts from the lumpy nature of plant-level investment. Given the prominent role of investment spending for shaping optimal monetary policy, this simplification could be problematic. Our analysis suggests, however, that this is not the case in the context of a New Keynesian model featuring lumpy investment à la Sveen and Weinke (2007).

JEL Codes: E22, E31, E32.

1. Introduction

New Keynesian (NK) theory often abstracts from capital accumulation (see, e.g., Galí 2015), and if capital accumulation is taken into account, then it is common practice to postulate convex adjustment costs of one type or another (see, e.g., Christiano, Eichenbaum, and Evans 2005 and Woodford 2005), which makes the model inconsistent with the observed lumpiness in plant-level investment (see, e.g., Thomas 2002, Khan and Thomas 2008). Given the prominent role of investment spending for shaping optimal monetary policy,[1]

*Thanks to seminar participants at the University of Konstanz as well as to co-editor Stephanie Schmitt-Grohé and two anonymous referees. Corresponding author: Lutz Weinke, Humboldt-Universität zu Berlin, School of Business and Economics, Institute of Economic Policy, Spandauer Straße 1, D-10099 Berlin, Germany, Tel.: +49 30 2093-5927, Fax: +49 30 2093-5934, E-mail: lutz.weinke@wiwi.hu-berlin.de.

[1]In the words of Schmitt-Grohé and Uribe (2007): "All the way from the work of Keynes (1936) and Hicks (1939) to that of Kydland and Prescott (1982) macroeconomic theories have emphasized investment dynamics as an important channel for the transmission of aggregate disturbances. It is therefore natural to expect that investment spending should play a role in shaping optimal monetary policy."

this simplification could be problematic. The reason is that lumpy investment is a potential source of significant heterogeneity across firms.[2] Hence, it might play an important role in shaping the optimal monetary policy response to shocks, the same way staggered prices generate inefficient price dispersion and thus provide a motive for inflation stabilization.

Our analysis shows, however, that the implications for the optimal design of monetary policy that obtain in the context of a conventional NK model depend very little on whether investment is lumpy, as in Sveen and Weinke (2007), or smooth, as in Woodford (2005). In Sveen and Weinke (2007) we assumed that both price-setting and investment decisions are conducted in a time-dependent fashion.[3] Up to a first-order approximation, the resulting framework has been shown to be observationally equivalent to the model in Woodford (2005) where he had postulated a convex capital adjustment cost at the firm level. In other words, the two alternative ways of modeling endogenous capital give rise to identical implications for aggregate dynamics. Note that in the field of investment theory, lumpiness is routinely generated by assuming stochastic non-convex capital adjustment costs (see, e.g., Thomas 2002, Khan and Thomas 2008). The recent contribution by Reiter, Sveen, and Weinke (2013) shows, however, that once that modeling choice is embedded into an otherwise conventional NK model, the latter implies large but very short-lived impacts on output and inflation, in a way that goes against the empirical evidence. In other words, it is well established, by now, that standard investment theory gives rise to counterfactual implications in the context of NK theory.[4] We therefore conduct our welfare analysis in the context of the model proposed in Sveen and Weinke (2007). The latter is the only lumpy investment model known so far that gives rise to an empirically relevant monetary transmission mechanism. Apart from the restrictions on capital accumulation, the models under consideration feature sticky prices and wages à la Calvo (1983). The co-existence of the two types of nominal

[2] In the context of the models presented in this paper, there is no distinction between a firm and a plant. We therefore use those terms interchangeably.

[3] The assumption of time-dependent investment was originally proposed by Kiyotaki and Moore (1997).

[4] The simple reason is the very large interest rate sensitivity of investment implied by that investment theory.

rigidities gives rise to a monetary policy tradeoff and consequently to an interesting monetary policy design problem, as discussed in Erceg, Henderson, and Levin (2000).

Based on Rotemberg and Woodford (1997) and Sveen and Weinke (2009), the present paper presents a second-order accurate expression for the level of welfare[5] in the context of a baseline model featuring time-dependent lumpy investment. The welfare criterion is the unconditional expectation of average household utility. For a benchmark model where firm-level capital is assumed to be smooth as in Woodford (2005), that welfare criterion is of the form calculated in Sveen and Weinke (2009). Our first main result shows that baseline and benchmark imply optimized interest rate rules which are similar in the following sense: if the optimized rule from one model is used in the other one, then the resulting additional welfare loss is negligible. This is an important extension of our earlier work in Sveen and Weinke (2009). In that paper we had shown that optimized interest rate rules, as implied by Woodford's (2005) modeling of endogenous capital, are very similar (in the same sense as above) to the ones that obtain under the assumption of a rental market for capital after an appropriate adjustment of the price stickiness in the latter model. The reason for the adjustment is that Woodford's (2005) modeling of investment implies some endogenous price stickiness, i.e., a price setter internalizes the consequences of a price-setting decision for the marginal costs over the expected lifetime of the chosen price. This effect is, however, absent in the rental market model, where the marginal cost is constant across all firms in the economy. This difference in price stickiness turns out to be the only first-order difference between Woodford's (2005) model and a rental market specification, as analyzed in Sveen and Weinke (2005). After making the two models equivalent up to the first order (by adjusting the price stickiness), the corresponding welfare implications become very similar. Much of the related literature on the optimal design of interest rate rules has adopted the assumption of a rental market for capital.[6] It is therefore interesting to observe that our first result in

[5]A meaningful welfare analysis can only be conducted based on a second-order approximation to the welfare criterion. See, e.g., Schmitt-Grohé and Uribe (2004).

[6]See, e.g., Schmitt-Grohé and Uribe (2007).

the present paper (when combined with the above-mentioned result in Sveen and Weinke 2009) lends some support to the view that those findings in the related literature appear to hinge very little on the rental market assumption. Second, we compare the welfare implications of each simple interest rate rule to the corresponding outcome under Ramsey optimal policy. This way we demonstrate that optimal policy is well approximated by optimized simple interest rate rules. This increases the practical relevance of our first result.

The remainder of the paper is organized as follows. Section 2 outlines the model structure. In section 3 the results are presented and section 4 concludes.

2. Model

We use a New Keynesian model with complete markets. Shocks to total factor productivity are assumed to be the only source of aggregate uncertainty. There is a continuum of households and a continuum of firms. Each household (firm) is the monopolistically competitive supplier of a differentiated type of labor (type of good) and we assume sticky wages (sticky prices) à la Calvo (1983), i.e., each household (firm) gets to reoptimize its wage (price) with a constant and exogenous probability. Capital accumulation is assumed to take place at the firm level, and the additional capital resulting from an investment decision becomes productive with a one-period delay. In our baseline model (*LI* for short) lumpiness in investment is modeled as in Sveen and Weinke (2007), i.e., there is a Bernoulli draw conditional on which a firm is allowed to invest, or not. In the benchmark model (*FS* for short) firm-level investment is (unrealistically) assumed to be smooth, as in Woodford (2005).[7] Since the details of the models have been discussed elsewhere (see Erceg, Henderson, and Levin 2000, Woodford 2005, and Sveen and Weinke 2007, 2009), we turn directly to the implied linearized equilibrium conditions. The only two equations that are different in *LI* and *FS* are, respectively, the inflation equation and the law of motion of

[7] *FS* stands for "firm specific." It should be noted, however, that capital in the baseline model is also firm specific. We refer to the baseline specification as being the lumpy investment model (*LI*) and call the benchmark a model of firm-specific capital since this was Woodford's (2005) original label for this type of model.

capital, as we are going to see. Relying on the insights in Rotemberg and Woodford (1997) and Sveen and Weinke (2009), we will compute our welfare criterion (in each model under consideration) up to the second order.

2.1 Some Linearized Equilibrium Conditions

In what follows we present our baseline model *LI* and make clear in which places it differs from the benchmark *FS*. We consider a linear approximation around a zero-inflation steady state. In what follows, variables are expressed in terms of log-deviations from their steady-state values except for the nominal interest rate, i_t, wage inflation, ω_t, and price inflation, π_t, which denote the levels of the respective variables. The consumption Euler equation reads

$$c_t = E_t c_{t+1} - \left(i_t - E_t \pi_{t+1} - \rho\right), \tag{1}$$

where c_t denotes aggregate consumption and E_t is the expectation operator conditional on information available through time t. Moreover, parameter ρ is the discount rate and the last equation also reflects our assumption of log consumption utility. Up to the first order, aggregate production is pinned down by

$$y_t = x_t + \alpha \; k_t + (1 - \alpha) \; n_t, \tag{2}$$

where y_t denotes aggregate output, parameter α indicates the capital share, aggregate capital is k_t, aggregate hours are n_t, and x_t represents an exogenous index of technology. The latter is assumed to follow an $AR(1)$ process $x_t = \rho_x \; x_{t-1} + \varepsilon_t$, with $\rho_x \in (0, 1)$ and ε_t denoting an iid shock with variance σ^2. The wage-inflation equation results from averaging and aggregating optimal wage-setting decisions on the part of households, as discussed in Erceg, Henderson, and Levin (2000). It takes the following simple form:

$$\omega_t = \beta \; E_t \omega_{t+1} + \lambda_\omega \; \left(mrs_t - rw_t\right), \tag{3}$$

where parameter $\beta \equiv 1/(1 + \rho)$ is the subjective discount factor, rw_t denotes the real wage, and $mrs_t \equiv c_t + \eta n_t$ measures the average marginal rate of substitution of consumption for leisure. In the latter definition parameter η indicates the inverse of the (aggregate)

Frisch labor supply elasticity. Finally, we have used the definition $\lambda_\omega \equiv \frac{(1-\beta\theta_w)(1-\theta_w)}{\theta_w}\frac{1}{1+\eta\varepsilon_N}$, where parameter θ_w denotes the probability that a household is not allowed to reoptimize its nominal wage in any given period, while parameter ε_N measures the elasticity of substitution between different types of labor. The price inflation equation takes the form

$$\pi_t = \beta\ E_t\pi_{t+1} + \lambda\ mc_t, \tag{4}$$

where $mc_t \equiv rw_t - (y_t - n_t)$ denotes the average real marginal cost. We have also used the definition $\lambda \equiv \frac{(1-\beta\theta)(1-\theta)}{\theta}\frac{1-\alpha}{1-\alpha+\alpha\varepsilon}\frac{1}{\xi}$, where parameter θ gives the probability that a firm does not get to reoptimize its price in any given period, while parameter ε denotes the elasticity of substitution between the differentiated goods. Finally, coefficient ξ is a function of the model's structural parameters that is computed numerically as in Sveen and Weinke (2007), using the method developed in Woodford (2005). The details of the respective computations of coefficient ξ depend on whether firm-level investment is lumpy, as in *LI*, or smooth as in *FS*. For future reference, it should be pointed out that Woodford's (2005) method relies on observing that a firm's price-setting rule can be approximated up to the first order by the following expression:

$$\widehat{p_t^*}(i) = \hat{p}_t^* - \tau_1\ \hat{k}_t(i),$$

with $\widehat{p_t^*}(i) \equiv p_t^*(i) - p_t$, $\hat{k}_t(i) \equiv k_t(i) - k_t$, and $\widehat{p_t^*} \equiv p_t^* - p_t$, where $p_t^*(i)$ is the price chosen by firm i in period t and p_t denotes the aggregate price level associated with the usual Dixit-Stiglitz aggregator. Firm i's relative to average capital stock is given by $\hat{k}_t(i)$. Moreover, $\hat{p}_t^*$ is the average relative price in the group of time-t price setters. Finally, coefficient τ_1 can be calculated by using Woodford's (2005) method. Here again the details of the respective computations of coefficient τ_1 depend on whether firm-level investment is lumpy, as in *LI*, or smooth as in *FS*. The law of motion of capital is obtained from averaging and aggregating optimal investment decisions on the part of firms. This implies

$$\begin{aligned}\Delta k_{t+1} = \beta\ E_t\Delta k_{t+2} + \frac{1}{\epsilon_\psi}\left[(1-\beta(1-\delta))\,E_t ms_{t+1}\right.\\ \left. - (i_t - E_t\pi_{t+1} - \rho)\right],\end{aligned} \tag{5}$$

where $ms_t \equiv rw_t - (k_t - n_t)$ represents the average real marginal return to capital. The latter is measured in terms of marginal savings in labor costs since firms are demand constrained in our model. Moreover, parameter δ is the rate of depreciation. Finally, coefficient ϵ_ψ measures the smoothness in aggregate capital accumulation. It is a function of the model's structural parameters that is evaluated as in Sveen and Weinke (2007), where we observe that a firm's investment rule can be approximated up to the first order by the following expression:

$$\widehat{k}^*_{t+1}(i) = \hat{k}^*_{t+1} - \tau_2\ \hat{p}_t(i),$$

with $\widehat{k}^*_{t+1}(i) \equiv k^*_{t+1}(i) - k_{t+1}$, $\widehat{k}^*_{t+1} \equiv k^*_{t+1} - k_{t+1}$, and $\hat{p}_t(i) \equiv p_t(i) - p_t$, where $k^*_{t+1}(i)$ is the capital stock chosen by firm i in period t. Moreover, k^*_{t+1} is the average capital stock in the group of time-t investors and $\hat{p}_t(i)$ is firm i's relative price in period t. Coefficient τ_2 can also be calculated by using Woodford's (2005) method. The value of parameter ϵ_ψ in (5) depends crucially on the probability θ_k with which a firm is not allowed to invest in any given period. Note that ϵ_ψ is simply the parameter measuring the convex capital adjustment cost in *FS*. Finally, we observe that the value of λ in (4) coincides with its counterpart in *FS* if parameter θ_k is chosen in such a way that the laws of motion of capital coincide in *LI* and *FS*. This is the reason why the two models are observationally equivalent, up to the first order.[8] The goods market clearing condition reads

$$y_t = \zeta\ c_t + \frac{1-\zeta}{\delta}\left[k_{t+1} - (1-\delta)\,k_t\right], \tag{6}$$

where $\zeta \equiv 1 - \frac{\delta\alpha}{\rho+\delta}$ denotes the steady-state consumption-to-output ratio. Notice that the frictionless desired markup, $\mu \equiv \frac{\varepsilon}{\varepsilon-1}$, does not enter the latter definition. The reason is our assumption of a wage subsidy that makes the steady state of our model efficient. This allows us to approximate our welfare criterion up to the second order relying on the first-order approximation that we have just considered. Next we discuss the welfare criterion that will be used

[8] For details, see Sveen and Weinke (2007).

to assess the desirability of alternative arrangements for the conduct of monetary policy.

2.2 Welfare

As in Erceg, Henderson, and Levin (2000), the policymaker's period welfare function is assumed to be the unweighted average of households' period utility

$$\mathcal{W}_t \equiv U\left(C_t\right) + \int_0^1 V\left(N_t\left(h\right)\right)dh = U\left(C_t\right) + E_h\left\{V\left(N_t\left(h\right)\right)\right\}. \quad (7)$$

The assumption of complete asset markets combined with a utility function that is separable in its two arguments, consumption and hours worked, implies that heterogeneity in hours worked (a consequence of wage dispersion) does not translate into consumption heterogeneity. This is reflected in (7), where $U\left(C_t\right)$ indicates consumption utility and $V\left(N_t\left(h\right)\right)$ denotes the disutility associated with household h's decision to supply $N_t\left(h\right)$ hours. The notation E_h is meant to indicate a cross-sectional expectation integrating over households. We also follow Erceg, Henderson, and Levin (2000) in expressing period welfare as a fraction of steady-state Pareto-optimal consumption, i.e., we consider $\frac{\mathcal{W}_t - \overline{\mathcal{W}}}{\overline{U}_C \overline{C}}$, where a bar indicates the steady-state value of the original variable and U_C is the marginal utility of consumption. The second-order approximation to period welfare is calculated based on the method by Rotemberg and Woodford (1997).[9] Our welfare criterion is the unconditional expectation of period welfare, which we write in the way proposed by Svensson (2000):

$$\widetilde{\mathcal{W}} \equiv \lim_{\beta \to 1} E_t\left\{(1-\beta)\sum_{k=0}^{\infty}\beta^k\left(\frac{\mathcal{W}_{t+k} - \overline{W}}{\overline{U}_C\overline{C}}\right)\right\}.$$

[9]The proof that the method of Rotemberg and Woodford (1997) can be applied to the problem at hand carries over from Edge's (2003) work to ours because the relevant steady-states properties of the two models are identical.

We derive the following expression for that welfare criterion in *LI*:

$$\widetilde{\mathcal{W}} \simeq E\left\{\Omega_1 y_t^2 + \Omega_2 c_t^2 + \Omega_3 k_t^2 + \Omega_4 n_t^2 + \Omega_5 \omega_t^2 + \Omega_6 \pi_t^2 + \Omega_7 (\Delta k_t)^2 + \Omega_8 z_t^2\right\}, \tag{8}$$

where the symbol $\simeq$ is meant to indicate that an approximation is accurate up to the second order. We have also used the definition $z_t \equiv \Delta k_t + \pi_t$, and coefficients Ω_1 to Ω_8 are functions of the structural parameters, as defined in the appendix. For *FS* the welfare criterion is calculated along the lines of Sveen and Weinke (2009). In order to derive (8), one needs to calculate the cross-sectional variances of prices and capital holdings at each point in time as well as their covariance. This is shown in the appendix. In a nutshell, we note that the above-mentioned linearized rules for price setting and for investment can be used to compute the relevant second moments with the accuracy that is needed for our second-order approximation to the welfare criterion.

3. Results

We consider a conventional family of monetary policy rules and analyze constrained optimal rules, i.e., we restrict attention to a particular subset of possible parameter values that parameterize the rule. It is useful to note that rational expectations equilibrium is locally unique (i.e., determinate) under the constrained optimal policies. The simple interest rate rules under consideration below are therefore also implementable. We first compare the welfare implications of constrained optimal simple and implementable interest rate rules in *LI* and *FS*. In a second step, we ask how well those optimal rules approximate the Ramsey optimal policy.

3.1 Baseline Parameter Values

In our quantitative analysis, the following values are assigned to the model parameters.[10] We consider a quarterly model. The capital share, α, is set to 0.36. The elasticity of substitution between goods,

[10]To solve the dynamic stochastic system of equations, we use Dynare (www.dynare.org). MATLAB code for our implementation of Woodford's (2005)

ε, takes the value 11. The rate of capital depreciation, δ, is assumed to be equal to 0.025. We set the elasticity of substitution between different types of labor, ε_N, equal to 6. Our choice for the Calvo price stickiness parameter, θ, is 0.75 and the value for the wage stickiness parameter, θ_w, is 0.75. This implies an average expected duration of price and wage contracts of one year. The coefficient of autocorrelation in the process of technology, ρ_x, is assumed to take the value 0.95. Those parameter values are justified in Erceg, Henderson, and Levin (2000), Sveen and Weinke (2007, 2009), and the references therein. Finally, we set the lumpiness parameter $\theta_k = 0.955$. The latter choice makes our model consistent with the evidence in Khan and Thomas (2008), according to which 18 percent of plants undertake lumpy investments in any given year. Conditional on our baseline choices for the remaining parameters, this results in $\epsilon_\psi = 6.359$, a value implying a plausible degree of smoothness in aggregate capital accumulation.[11]

3.2 The Welfare Consequences of Simple Implementable Rules

We consider a family of simple interest rate rules of the form

$$r_t = \rho + \tau_r \ (r_{t-1} - \rho) + \tau_s \left[\tau_\omega \omega_t + (1 - \tau_\omega) \pi_t\right] + \tau_y \ \tilde{y}_t, \tag{9}$$

where parameter τ_s measures the overall responsiveness of the nominal interest rate to changes in inflation, whereas τ_ω is the relative weight put on wage inflation. The weight on price inflation is therefore given by $(1 - \tau_\omega)$. Parameter τ_r denotes the interest rate smoothing coefficient and parameter τ_y indicates the quantitative importance attached to a measure of real economic activity that is taken to be the output gap, $\tilde{y}_t$. To be precise, $\tilde{y}_t \equiv y_t - y_t^{nat}$, where y_t^{nat} is natural output that is computed à la Neiss and Nelson (2003), i.e., the equilibrium output that would obtain in the absence

method is available upon request. We also thank Junior Maih for code that has been used in the computation of the optimized coefficients of the interest rate rules under consideration.

[11] As we have mentioned already, ϵ_ψ is simply the parameter measuring the convex capital adjustment cost in Woodford (2005).

of nominal frictions. Only positive parameter values are considered, and parameter τ_ω is required to be less than or equal to one.

We follow Erceg, Henderson, and Levin (2000) and consider the business-cycle cost of nominal rigidities. For each model under consideration, we report the optimized coefficients entering the interest rate rule as well as the corresponding welfare measure. Specifically, we compute our welfare measure in each model under the assumption of flexible prices and wages (i.e., $\widetilde{\mathcal{W}}^{nat}$) and subtract the resulting expression from the value in the corresponding model with the nominal rigidities being present (i.e., $\widetilde{\mathcal{W}}$). As they do, we divide this difference by the productivity innovation variance, σ^2, and refer to the resulting number as being a welfare loss. Let us give a concrete example for the interpretation of the welfare loss numbers in our tables. Suppose the productivity innovation variance is 0.00315^2. Under a standard Taylor (1993) rule for the conduct of monetary policy (i.e., $\tau_r = 0.7$, $\tau_s = 1.5$, $\tau_\omega = 0$, and $\tau_y = 0.125$), this choice implies an output variance of 0.016^2, which is well in line with the variance of the cyclical component of output in the U.S. economy, at least for the period starting in the early eighties until the onset of the 2007 financial crisis. Then, a welfare loss of -10 means that the representative household would be willing to give up $10 \times 0.00315^2 \times 100 = 0.01$ percentage points of steady-state (Pareto-optimal) consumption in order to avoid the business-cycle cost associated with the presence of the nominal rigidities.

3.2.1 Wage–Price Rule

We now compare optimized interest rate rules under *LI* and *FS*. Our first set of results regards interest rate rules according to which the nominal interest rate is only adjusted in response to changes in price and wage inflation, i.e., parameter τ_y is set to zero. We also state the welfare loss implied by *LI* if the optimized rule from *FS* is used in that model. The results are shown in table 1.

The welfare losses under the respective optimized rules are very similar in *LI* and *FS*, and the differences between the corresponding optimized parameter values, as implied by the two models, are also small. Moreover, the additional welfare loss is tiny if the optimized rule from *FS* is used in *LI* (compared with the outcome that obtains under the optimized rule in *LI*). In other words, a central banker

Table 1. Wage–Price Rule

Parameter	FS	LI
τ_r	1.0352	1.0353
τ_s	0.8399	0.5837
τ_ω	0.3416	0.3295
Welfare Loss $\left(\frac{\widetilde{\mathcal{W}}-\widetilde{\mathcal{W}}^{nat}}{\sigma^2}\right)$	-9.1104 (-9.1342^a)	-9.1264

[a] Welfare loss in the lumpy investment model with *FS* rule.

would only make a tiny mistake by abstracting from lumpiness in investment when optimizing a wage–price interest rate rule. What is the intuition? There is only a little interaction between price setting and investment in the two models under consideration. In fact, an investment decision is almost unaffected by the current price of a firm. The reason is that an investor takes rationally into account the future price changes that are expected to take place over the lifetime of the chosen capital stock. This in turn is so as long as investment decisions are forward-looking enough (regardless of the particular type of restriction on capital accumulation). Put differently, the additional heterogeneity among firms that is implied by the lumpiness in investment does not matter much for the business-cycle cost of the nominal rigidities.

To some extent, the results in table 1 are also surprising. To see this, note a subtle first-order difference between the versions of *LI* and *FS* that obtain under the assumption of flexible prices and wages (and keep in mind that those versions of the two models enter our calculations when isolating the business-cycle cost of the nominal rigidities). The reason is that ϵ_ψ is simply a parameter in the law of motion of capital, as implied by *FS*. On the other hand, ϵ_ψ is a coefficient in the context of *LI* whose value also depends (among other parameter values) on the degree of price stickiness, as analyzed in Sveen and Weinke (2007). This difference turns out, however, to be inconsequential for the results shown in table 1 according to which the two models have very similar implications for the optimal design of wage–price rules. It is shown next that the similarity between the welfare results in *LI* and *FS* is not specific to the type of interest rate rule that we had considered so far.

3.2.2 Taylor-Type Rule

Once again, we compare welfare implications in *LI* and *FS*. The structure of the comparison is the same as in the preceding subsection, but we now consider interest rate rules according to which the nominal interest rate is only adjusted in response to changes in price inflation and the output gap, i.e., parameter τ_ω is set to zero. Also in this case, our analysis brings out a clear message. This is shown in table 2.

There are only very small differences between the welfare losses under the optimized rules in *LI* and *FS*, and the same is true for the respective coefficients parameterizing those rules. Moreover, the additional welfare loss is once again tiny if the optimized rule from *FS* is used in *LI* (compared with the outcome that obtains under the optimized rule in *LI*). We have already mentioned that the law of motion of natural capital is different in the two models because price stickiness affects the smoothness of aggregate capital accumulation in *LI*, but not in *FS*. This observation is particularly relevant for the analysis of Taylor-type interest rate rules because natural output is used in the construction of the output gap. However, as our results clearly show, the quantitative importance of this effect is also very limited in this context. Next we analyze how well the simple and implementable interest rate rules considered so far approximate Ramsey-optimal policy.

3.3 Ramsey-Optimal Policy

We now compute Ramsey-optimal welfare losses in *LI* and *FS*.[12] This is shown in table 3.

In each model the welfare loss under Ramsey-optimal policy is well approximated by the corresponding outcomes that are associated with the optimized interest rate rules considered in the preceding subsection. Taken together, the present paper therefore conveys a positive message. We develop a New Keynesian model featuring lumpy investment using the formalism of a time-dependent rule à la Sveen and Weinke (2007). (As we discussed in the introduction, no alternative lumpy investment model has been proposed so far in the

[12] To compute Ramsey optimal policy, we use Dynare (www.dynare.org).

Table 2. Taylor-Type Rule with Neiss and Nelson Output Gap

Parameter	FS	LI
τ_r	1.0423	0.9911
τ_s	0.0365	0.0093
τ_y	0.0283	0.0074
Welfare Loss $\left(\frac{\widetilde{\mathcal{W}}-\widetilde{\mathcal{W}}^{nat}}{\sigma^2}\right)$	-9.1914 (-9.4744^a)	-9.1639

[a]Welfare loss in the lumpy investment model with FS rule.

Table 3. Ramsey-Optimal Policy

	FS	LI
Welfare Loss $\left(\frac{\widetilde{\mathcal{W}}-\widetilde{\mathcal{W}}^{nat}}{\sigma^2}\right)$	-8.9311	-8.9157

literature that would give rise to an empirically relevant monetary transmission mechanism.) In the context of the resulting framework, it turns out that the conclusions for the desirability of alternative arrangements for the conduct of monetary policy are strikingly similar regardless of whether or not lumpiness in investment is taken into account.

3.4 Sensitivity Analysis

Let us challenge the results obtained so far. Our first sensitivity analysis regards the notion of the output gap. In the context of a model featuring endogenous capital accumulation, Woodford (2003, chapter 5) proposes to refine that notion in the following way. He uses the equilibrium output that would obtain if the nominal rigidities were absent and expected to be absent in the future but taking as given the capital stock resulting from optimizing investment behavior in the past in an environment *with* the nominal rigidities present. Woodford argues that this measure of natural output is more closely related to equilibrium determination than the alternative measure

Table 4. Taylor-Type Rule with Woodford Output Gap

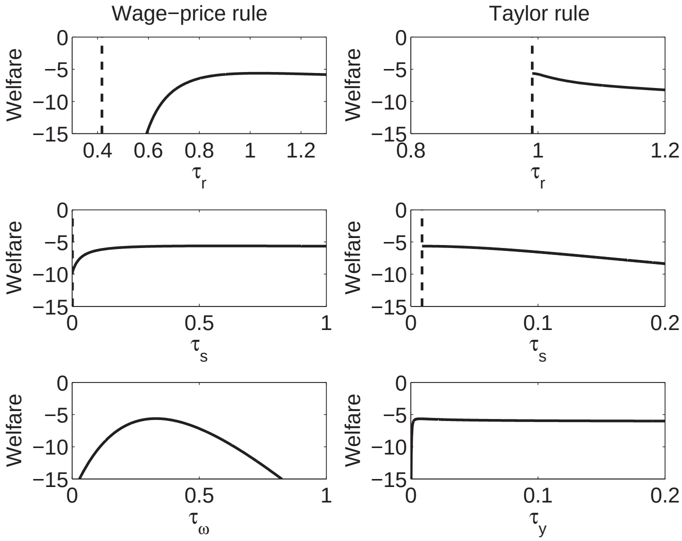
Parameter	**FS**	**LI**
τ_r	1.0305	1.0080
τ_π	0.0545	0
τ_y	0.0227	0.0343
Welfare Loss $\left(\frac{\widetilde{\mathcal{W}}-\widetilde{\mathcal{W}}^{nat}}{\sigma^2}\right)$	–9.1935 (–9.3054[a])	–9.0641

[a]Welfare loss in the lumpy investment model with *FS* rule.

that has been proposed by Neiss and Nelson (2003) and that we have used so far in our analysis. We have stated already that under their definition natural output is the equilibrium output that would obtain if nominal rigidities were absent. But this means that they are not only currently absent and expected to be absent in the future but that they had also been absent *in the past*. This distinction turns out, however, to be of negligible importance for the results in the present paper. This is shown in table 4.

Our second sensitivity analysis is concerned with the following question: What are the additional welfare losses implied by deviations from the optimized interest rate rules that have been analyzed so far in the context of *LI*? Let us therefore reconsider the general interest rate rule stated in equation (9). The left-hand side of figure 1 shows by how much the welfare loss increases if the policy parameters deviate from their optimal values in the case of the wage–price rule considered above (i.e., the specification where parameter τ_y is set to zero). The right-hand side of figure 1 shows the corresponding analysis for the case of the Taylor-type rule (i.e., the specification where parameter τ_ω is set to zero). Each panel of the figure illustrates the welfare loss associated with a change in one policy parameter at a time while holding the remaining parameters constant at their optimal values. In each case, the figure also shows the corresponding indeterminacy region in the parameter space. More concretely, dashed lines indicate the critical parameter values in the sense that rational expectations equilibrium is locally unique (or determinate) for all parameter values that are larger than those critical values. Finally, let us mention that each panel in the figure shows the same

Figure 1. Robustness Analysis of Welfare Losses Associated with Simple Implementable Interest Rate Rules in the Context of *LI*



range of welfare losses. That range is chosen arbitrarily, but we believe that our particular choice allows us to highlight the most relevant economic effects. In some cases (more concretely, for the wage–price rule and its parameters τ_r and τ_ω) parameter values that give rise to a determinate equilibrium imply welfare losses outside that range. Those losses are therefore not shown in the corresponding panels of figure 1.

From an economic point of view, our sensitivity analysis with respect to the output gap response in the Taylor-type rule deserves special attention. It is shown that small deviations from the optimal rule result in negligible welfare losses. Only to the extent that the central bank attaches very little importance to the output gap in setting the nominal interest rate does the resulting rule imply a large welfare loss with respect to the optimal policy.[13]

[13]Let us also mention here that the conventional wisdom according to which responding to the level of output in an interest rate rule is costly from a welfare point of view (see, e.g., Schmitt-Grohé and Uribe 2007) remains valid in the context of the models analyzed in the present paper.

4. Conclusion

We augment the model in Erceg, Henderson, and Levin (2000) with a standard textbook treatment of endogenous capital accumulation in the context of New Keynesian theory, namely the firm-specific investment decision considered in Woodford (2005). The latter model abstracts, however, from lumpiness in investment, which is an uncontroversial empirical regularity. The present paper shows that the implications of Woodford's (2005) modeling of capital accumulation turn out to be similar to the ones that emerge in the presence of a lumpy investment decision of the form considered in Sveen and Weinke (2007). In fact, under both alternative assumptions on capital accumulation, we find that simple implementable interest rate rules generate welfare results that are similar. Moreover, the additional welfare loss associated with using the optimized rule from one model in the other one is tiny. Finally, optimized simple interest rate rules are shown to approximate very well the corresponding outcomes under Ramsey-optimal policy.

We had already mentioned that there is no consensus on the way in which endogenous capital accumulation should be modeled in the context of monetary models. Clearly, this issue is high on our research agenda, and, needless to say, alternative ways of modeling lumpiness in investment would change not only the implied monetary transmission mechanism but also the model's welfare implications. Another caveat regards the assumed stochastic driving force of the models under consideration. In each case we restrict attention to stationary exogenous shocks to total factor productivity. This is in line with the assumption in Erceg, Henderson, and Levin (2000) on which we build our analysis. Given the uncontroversial empirical support for the importance of demand factors for aggregate fluctuations, it would however also be interesting to assess the potential role of alternative economic shocks for the results obtained in the present paper. Those caveats notwithstanding, our results demonstrate that time-dependent lumpy investment in itself does not lead to any important changes in the welfare relevant results with respect to standard textbook treatments of endogenous capital in the New Keynesian theory.

Appendix. Welfare with Lumpy Investment

Throughout the appendix we use the notation and the definitions that are introduced in the text. We approximate our utility-based welfare criterion up to the second order. In what follows, we make frequent use of two rules,

$$\frac{A_t - \overline{A}}{\overline{A}} \simeq a_t + \frac{1}{2}a_t^2, \tag{10}$$

where $\simeq$ is meant to indicate that an approximation is accurate up to the second order. We have also used the definition $a_t \equiv \ln\left(\frac{A_t}{\overline{A}}\right)$.

Moreover, if $A_t = \left(\int\limits_0^1 A_t(i)^\gamma \, di\right)^{\frac{1}{\gamma}}$, then

$$a_t \simeq E_i a_t(i) + \frac{1}{2}\gamma Var_i a_t(i), \tag{11}$$

with Var_i indicating a cross-sectional variance. As we have already mentioned in the text, the policymaker's period welfare function reads

$$\mathcal{W}_t \equiv U(C_t) - \int_0^1 V(N_t(h))\, dh = U(C_t) - E_h\{V(N_t(h))\}. \tag{12}$$

It follows that

$$\mathcal{W}_t \simeq \overline{\mathcal{W}} + \overline{U}_C\overline{C}\left(c_t + \frac{1}{2}c_t^2\right) - \overline{V}_N\overline{N}E_h\left\{n_t(h) + \frac{1}{2}n_t(h)^2\right\}$$
$$+ \frac{1}{2}\overline{U}_{CC}\overline{C}^2 c_t^2 - \frac{1}{2}\overline{V}_{NN}\overline{N}^2 E_h\left\{n_t(h)^2\right\}.$$

The Pareto optimality of the steady state implies $\overline{V}_N\overline{N} = U_C\overline{C}\frac{1-\alpha}{\zeta}$. We therefore have

$$\frac{\mathcal{W}_t - \overline{\mathcal{W}}}{\overline{U}_C\overline{C}} \simeq \left(c_t + \frac{1}{2}c_t^2\right) - \frac{1}{2}c_t^2 - \frac{1-\alpha}{\zeta}E_h\{n_t(h)\}$$
$$- \frac{1}{2}\frac{1-\alpha}{\zeta}(1+\phi)E_h\left\{n_t(h)^2\right\}. \tag{13}$$

Next we show how the linear terms in consumption and employment in the last expression can be approximated up to the second order. We start by analyzing the consumption portion of welfare. To this end we invoke the resource constraint.

The Consumption Portion of Welfare

The resource constraint reads

$$Y_t = C_t + K_{t+1} - (1-\delta)\, K_t,$$

where $K_t \equiv \int_0^1 K_t(i)\, di$ and $K_t(i)$ is the amount of composite goods used in firm i's production. It follows that

$$\begin{aligned} c_t + \frac{1}{2}c_t^2 \simeq \frac{1}{\zeta}\left(y_t + \frac{1}{2}y_t^2\right) - \frac{1-\zeta}{\zeta}\frac{1}{\delta}\left(k_{t+1} + \frac{1}{2}k_{t+1}^2\right) \\ + \frac{1-\zeta}{\zeta}\frac{1-\delta}{\delta}\left(k_t + \frac{1}{2}k_t^2\right). \end{aligned} \tag{14}$$

The Labor Portion of Welfare

Aggregate labor supply is given by $N_t \equiv \left(\int_0^1 N_t(i)^{\frac{\epsilon_N - 1}{\epsilon_N}}\, di\right)^{\frac{\epsilon_N}{\epsilon_N - 1}}$. Using the second rule, we can write

$$n_t \simeq E_h n_t(h) + \frac{1}{2}\frac{\epsilon_N - 1}{\epsilon_N} Var_h n_t(h).$$

Aggregate labor demand reads

$$L_t \equiv \int_0^1 L_t(i)\, di = \int_0^1 \left(\frac{Y_t(i)}{X_t K_t(i)^{\alpha}}\right)^{\frac{1}{1-\alpha}} di = B_t^{\frac{1}{1-\alpha}},$$

where $B_t \equiv \left(\int_0^1 B_t(i)^{\frac{1}{1-\alpha}}\, di\right)^{1-\alpha}$ and $B_t(i) \equiv \frac{Y_t(i)}{X_t K_t(i)^{\alpha}}$. Clearing of the labor market implies that $N_t = L_t$. We therefore have

$$n_t = \frac{1}{1-\alpha} b_t.$$

Invoking the second rule, we obtain

$$b_t \simeq E_i b_t(i) + \frac{1}{2}\frac{1}{1-\alpha} Var_i b_t(i),$$

and we also note that

$$b_t(i) = y_t(i) - x_t - \alpha\ k_t(i).$$

The last result implies

$$E_i b_t(i) = E_i y_t(i) - x_t - \alpha E_i k_t(i),$$
$$Var_i b_t(i) = Var_i y_t(i) + \alpha^2 \kappa_t - 2\alpha Cov_i(y_t(i), k_t(i)),$$

with Var_i denoting the cross-sectional variance and Cov_i the covariance operator. We can therefore write

$$\begin{aligned} b_t &\simeq E_i y_t(i) - x_t - \alpha E_i k_t(i) \\ &\quad + \frac{1}{2}\frac{1}{1-\alpha}\left[Var_i y_t(i) + \alpha^2 \kappa_t - 2\alpha Cov_i(y_t(i), k_t(i))\right]. \end{aligned}$$

Let us also observe that

$$E_i k_t(i) \simeq k_t - \frac{1}{2} Var_i k_t(i),$$
$$E_i y_t(i) \simeq y_t - \frac{1}{2}\frac{\varepsilon - 1}{\varepsilon} Var_i y_t(i).$$

Combining the last three results, we arrive at

$$\begin{aligned} b_t &\simeq y_t - x_t - \alpha\ k_t + \frac{1}{2}\frac{\alpha}{1-\alpha} Var_i k_t(i) + \frac{1}{2}\frac{1}{\varepsilon}\left(\frac{1-\alpha+\alpha\varepsilon}{1-\alpha}\right) Var_i y_t(i) \\ &\quad - \frac{\alpha}{1-\alpha} Cov_i(y_t(i), k_t(i)). \end{aligned}$$

It follows that

$$\begin{aligned} E_h n_t(h) \simeq{} & \frac{1}{1-\alpha}(y_t - x_t - \alpha k_t) + \frac{1}{2}\frac{\alpha}{(1-\alpha)^2} Var_i k_t(i) \\ & + \frac{1}{2}\frac{1}{\varepsilon}\frac{1-\alpha+\alpha\varepsilon}{(1-\alpha)^2} Var_i y_t(i) \\ & - \frac{\alpha}{(1-\alpha)^2} Cov_i(y_t(i), k_t(i)) - \frac{1}{2}\frac{\varepsilon_N - 1}{\varepsilon_N} Var_h n_t(h). \end{aligned}$$

Using the demand functions for goods and labor services, we obtain

$$Var_i y_t\left(i\right) \simeq \varepsilon^2 \Delta_t,$$
$$Cov_i\left(y_t\left(i\right), k_t\left(i\right)\right) \simeq -\varepsilon \psi_t,$$
$$Var_h n_t\left(h\right) \simeq \varepsilon_N^2 \lambda_t,$$

where $\Delta_t \equiv Var_i \widehat{p}_t\left(i\right)$, $\psi_t \equiv Cov_i\left(\widehat{p}_t\left(i\right), k_t\left(i\right)\right)$, and $\lambda_t \equiv Var_h \widehat{w}_t\left(h\right)$, with $\widehat{w}_t\left(h\right) \equiv w_t\left(h\right) - w_t$. In the latter definition, $w_t\left(h\right)$ is household h's nominal wage and w_t denotes the aggregate wage level associated with the corresponding Dixit-Stiglitz aggregator. We therefore have

$$\begin{aligned} E_h n_t\left(h\right) \simeq {} & \frac{1}{1-\alpha}\left(y_t - x_t - \alpha k_t\right) + \frac{1}{2}\frac{1-\alpha+\alpha\varepsilon}{\left(1-\alpha\right)^2}\varepsilon\ \Delta_t \\ & + \frac{1}{2}\frac{\alpha}{\left(1-\alpha\right)^2}\ \kappa_t + \frac{\alpha\varepsilon}{\left(1-\alpha\right)^2}\ \psi_t - \frac{1}{2}\left(\varepsilon_N - 1\right)\varepsilon_N\ \lambda_t, \end{aligned} \tag{15}$$

with $\kappa_t \equiv Var_i \hat{k}_t\left(i\right)$. Finally, we note that $E_h\left\{n_t\left(h\right)^2\right\}$ can be written as

$$E_h\left\{n_t\left(h\right)^2\right\} \simeq \varepsilon_N^2\ \lambda_t + n_t^2. \tag{16}$$

The Welfare Function

Combining (13), (14), (15), and (16), we arrive at the following approximation to period welfare:

$$\begin{aligned} \frac{\mathcal{W}_t - \overline{\mathcal{W}}}{\overline{U}_C \overline{C}} \simeq {} & \frac{1}{\zeta}\ x_t - \frac{1}{\zeta}\frac{\alpha}{\rho+\delta}\left[k_{t+1} + \left(1+\rho\right)k_t\right] + \frac{1}{2}\frac{1}{\zeta}\ y_t^2 - \frac{1}{2}\ c_t^2 \\ & - \frac{1}{2}\frac{1-\zeta}{\zeta}\frac{1}{\delta}\ k_{t+1}^2 + \frac{1}{2}\frac{1-\zeta}{\zeta}\frac{1-\delta}{\delta}\ k_t^2 \\ & - \frac{1}{2}\frac{1-\alpha}{\zeta}\left(1+\phi\right)\ n_t^2 - \frac{1}{2}\frac{\varepsilon}{\zeta}\frac{1-\alpha+\alpha\varepsilon}{1-\alpha}\ \Delta_t \\ & - \frac{1}{2}\frac{1}{\zeta}\frac{\alpha}{1-\alpha}\ \kappa_t - \frac{1}{2}\frac{\left(1-\alpha\right)\varepsilon_N}{\zeta}\left(1+\phi\varepsilon_N\right)\ \lambda_t - \frac{1}{\zeta}\frac{\alpha\varepsilon}{1-\alpha}\ \psi_t. \end{aligned} \tag{17}$$

The first linear term in the last expression is proportional to the level of aggregate (log) technology, x_t, which is exogenous. The remaining linear terms are proportional to current and next period's aggregate capital, k_t and k_{t+1}. Next we compute $\frac{1-\beta}{\overline{U}_C \overline{C}} E_t \left\{ \sum_{k=0}^{\infty} \beta^k \left(\mathcal{W}_{t+k} - \overline{\mathcal{W}} \right) \right\}$, which allows us to invoke a result by Edge (2003). As in her model, the terms in aggregate capital cancel except for the initial one. Following the lead of Svensson (2000), we consider the limit for $\beta \to 1$. This allows us to abstract from initial conditions.[14]

$$\begin{aligned} \widetilde{\mathcal{W}}^{LI} \simeq\ & \Omega_1\ E\left\{y_t^2\right\} + \Omega_2\ E\left\{c_t^2\right\} + \Omega_3\ E\left\{k_t^2\right\} + \Omega_4\ E\left\{n_t^2\right\} \\ & + \Omega_\lambda\ E\left\{\lambda_t\right\} + \Omega_\Delta\ E\left\{\Delta_t\right\} + \Omega_\kappa\ E\left\{\kappa_t\right\} + \Omega_\psi\ E\left\{\psi_t\right\}, \end{aligned} \tag{18}$$

with

$$\begin{gathered} \Omega_1 \equiv \frac{1}{2}\frac{1}{\zeta},\ \Omega_2 \equiv -\frac{1}{2},\ \Omega_3 \equiv -\frac{1}{2}\frac{1-\zeta}{\zeta}, \\ \Omega_4 \equiv -\frac{1}{2}\frac{1-\alpha}{\zeta}\left(1+\phi\right),\ \Omega_\lambda \equiv -\frac{1}{2}\frac{\left(1-\alpha\right)\varepsilon_N}{\zeta}\left(1+\phi\varepsilon_N\right), \\ \Omega_\Delta \equiv -\frac{1}{2}\frac{\varepsilon}{\zeta}\frac{1-\alpha+\alpha\varepsilon}{1-\alpha},\ \Omega_\kappa \equiv -\frac{1}{2}\frac{1}{\zeta}\frac{\alpha}{1-\alpha},\ \Omega_\psi \equiv -\frac{1}{\zeta}\frac{\alpha\varepsilon}{1-\alpha}. \end{gathered}$$

Next we derive recursive formulations for the cross-sectional variances of wages, prices, and capital holdings as well as the welfare relevant covariance of prices and capital holdings. As in Erceg, Henderson, and Levin (2000), the cross-sectional variance of wages, λ_t, can be written in the following way:

$$\lambda_t \simeq \theta_w\ \lambda_{t-1} + \frac{\theta_w}{1-\theta_w}\ \left(\omega_t\right)^2,$$

and the unconditional expected value is

$$E\left\{\lambda_t\right\} \simeq \frac{\theta_w}{\left(1-\theta_w\right)^2} E\left\{\omega_t^2\right\}.$$

[14]For a formal proof, see Dennis (2007).

Next we derive recursive formulations for Δ_t, κ_t, and ψ_t. To this end we start by invoking the pricing rule and the capital accumulation rule,

$$\begin{aligned}
\widehat{p}_t^*(i) &\overset{FO}{\simeq} \widehat{p}_t^* - \tau_1 \widehat{k}_t(i),\\
\ln P_t^*(i) &\overset{FO}{\simeq} \ln P_t^* - \tau_1 (k_t(i) - k_t),\\
k_{t+1}^*(i) &\overset{FO}{\simeq} k_{t+1}^* - \tau_2 (\ln P_t(i) - \ln P_t),
\end{aligned}$$

where $\overset{FO}{\simeq}$ is used to indicate that the approximation holds up to the first order. Moreover, we define

$$\widetilde{p}_t \equiv E_i \ln P_t(i),$$

and, in an analogous manner, $\widetilde{p}_t^*$. We should note that $\widetilde{p}_t^* \overset{FO}{\simeq} \ln P_t^*$. For future reference, we first derive the following two expressions:

$$\begin{aligned}
\widetilde{p}_t - \widetilde{p}_{t-1} = E_i[\ln P_t(i) - \widetilde{p}_{t-1}] &= \theta_p E_i[\ln P_{t-1}(i) - \widetilde{p}_{t-1}]\\
&\quad + (1-\theta_p) E_i[\ln P_t^*(i) - \widetilde{p}_{t-1}],\\
&\overset{FO}{\simeq} (1-\theta_p)(\widetilde{p}_t^* - \widetilde{p}_{t-1}),\\
k_{t+1} - k_t = E_i[k_{t+1}(i) - k_t] &= \theta_k E_i[k_t(i) - k_t]\\
&\quad + (1-\theta_k) E_i\left[k_{t+1}^*(i) - k_t\right],\\
&\overset{FO}{\simeq} (1-\theta_k)\left(k_{t+1}^* - k_t\right).
\end{aligned}$$

We rewrite the price dispersion term, Δ_t, in the way proposed by Woodford (2003, chapter 6),

$$\begin{aligned}
\Delta_t &= E_i\left[(\ln P_t(i) - \widetilde{p}_{t-1})^2\right] - (\widetilde{p}_t - \widetilde{p}_{t-1})^2,\\
&= \theta_p E_i\left[(\ln P_{t-1}(i) - \widetilde{p}_{t-1})^2\right]\\
&\quad + (1-\theta_p) E_i\left[(\ln P_t^*(i) - \widetilde{p}_{t-1})^2\right] - (\widetilde{p}_t - \widetilde{p}_{t-1})^2.
\end{aligned}$$

After invoking the price-setting rule, we obtain the following recursive formulation for the measure of price dispersion Δ_t:

$$\begin{aligned}
\Delta_t &\simeq \theta_p\ \Delta_{t-1} + (1-\theta_p)\, E_i \left[\left(\widetilde{p}_t^* - \tau_1 \left(k_t\,(i) - k_t \right) - \widetilde{p}_{t-1} \right)^2 \right] \\
&\quad - \left(\widetilde{p}_t - \widetilde{p}_{t-1} \right)^2 \\
&\simeq \theta_p\ \Delta_{t-1} \\
&\quad + (1-\theta_p)\, E_i \left[\left(\frac{1}{1-\theta_p} \left(\widetilde{p}_t - \widetilde{p}_{t-1} \right) - \tau_1 \left(k_t\,(i) - k_t \right) \right)^2 \right] \\
&\quad - \left(\widetilde{p}_t - \widetilde{p}_{t-1} \right)^2 \\
&\simeq \theta_p\ \Delta_{t-1} + (1-\theta_p)\, \tau_1^2 Var_i k_t\,(i) + \left(\frac{1}{1-\theta_p} - 1 \right) \left(\widetilde{p}_t - \widetilde{p}_{t-1} \right)^2 \\
&\simeq \theta_p\ \Delta_{t-1} + (1-\theta_p)\, \tau_1^2 \kappa_t + \frac{\theta_p}{1-\theta_p} \pi_t^2 .
\end{aligned}$$

For the cross-sectional variance of capital holdings, κ_{t+1}, we obtain

$$\begin{aligned}
\kappa_{t+1} &= E_i \left[\left(k_{t+1}\,(i) - k_t \right)^2 \right] - \left(E_i k_{t+1}\,(i) - k_t \right)^2 \\
&= \theta_k E_i \left[\left(k_t\,(i) - k_t \right)^2 \right] + (1-\theta_k)\, E_i \left[\left(k_{t+1}^*\,(i) - k_t \right)^2 \right] \\
&\quad - \left(k_{t+1} - k_t \right)^2 \\
&\simeq \theta_k\ \kappa_t + (1-\theta_k)\, E_i \left[\left(k_{t+1}^* - k_t - \tau_2 \left(\ln P_t\,(i) - \ln P_t \right) \right)^2 \right] \\
&\quad - \left(k_{t+1} - k_t \right)^2 \\
&\quad - (1-\theta_k) \left[\left(k_{t+1}^* - k_t - \tau_2 \left(E_i \ln P_t\,(i) - \ln P_t \right) \right)^2 \right] \\
&\quad + \frac{1}{1-\theta_k} \left(k_{t+1} - k_t \right)^2 \\
&\simeq \theta_k\ \kappa_t + (1-\theta_k)\, \tau_2^2 \Delta_t + \frac{\theta_k}{1-\theta_k} \left(k_{t+1} - k_t \right)^2 .
\end{aligned}$$

Finally, the covariance between prices and capital holdings, ψ_t, takes the following form:

$$
\begin{aligned}
\psi_t &= E_i\left[\left(k_t\left(i\right)-k_{t-1}\right)\left(\ln P_t\left(i\right)-\ln P_{t-1}\right)\right] \\
&\quad -\left[\left(E_i k_t\left(i\right)-k_{t-1}\right)\left(E_i \ln P_t\left(i\right)-\ln P_{t-1}\right)\right]. \\
&= \theta_p\theta_k E_i\left[\left(k_{t-1}\left(i\right)-k_{t-1}\right)\left(\ln P_{t-1}\left(i\right)-\ln P_{t-1}\right)\right] \\
&\quad +\ \theta_p\left(1-\theta_k\right)E_i\left[\left(k_t^*\left(i\right)-k_{t-1}\right)\left(\ln P_{t-1}\left(i\right)-\ln P_{t-1}\right)\right] \\
&\quad +\left(1-\theta_p\right)\theta_k E_i\left[\left(k_{t-1}\left(i\right)-k_{t-1}\right)\left(\ln P_t^*\left(i\right)-\ln P_{t-1}\right)\right] \\
&\quad +\left(1-\theta_p\right)\left(1-\theta_k\right)E_i\left[\left(k_t^*\left(i\right)-k_{t-1}\right)\left(\ln P_t^*\left(i\right)-\ln P_{t-1}\right)\right] \\
&\quad -\left(k_t-k_{t-1}\right)\pi_t \\
&\simeq \theta_p\theta_k\psi_{t-1}+\theta_p\left(1-\theta_k\right)E_i\left[\left(k_t^*-k_{t-1}\right.\right. \\
&\quad \left.\left.-\tau_2\left(\ln P_{t-1}\left(i\right)-\ln P_{t-1}\right)\right)\left(\ln P_{t-1}\left(i\right)-\ln P_{t-1}\right)\right] \\
&\quad +\left(1-\theta_p\right)\theta_k E_i\left[\left(k_{t-1}\left(i\right)-k_{t-1}\right)\right. \\
&\quad \left.\times\left(\ln P_t^*-\tau_1\left(k_{t-1}\left(i\right)-k_{t-1}+k_{t-1}-k_t\right)-\ln P_{t-1}\right)\right] \\
&\quad +\left(1-\theta_p\right)\left(1-\theta_k\right)E_i\left[\left(k_t^*-k_{t-1}-\tau_2\left(\ln P_{t-1}\left(i\right)-\ln P_{t-1}\right)\right)\right. \\
&\quad \left.\times\ \left(\ln P_t^*-\ln P_{t-1}-\tau_1\left(k_t^*-k_t\right)+\tau_1\tau_2\left(\ln P_{t-1}\left(i\right)-\ln P_{t-1}\right)\right)\right] \\
&\quad -\left(k_t-k_{t-1}\right)\pi_t \\
&\simeq \theta_p\theta_k\ \psi_{t-1}-\tau_2\left(1-\theta_k\right)\left[\theta_p+\tau_1\tau_2\left(1-\theta_p\right)\right]\Delta_{t-1} \\
&\quad -\tau_1\left(1-\theta_p\right)\theta_k\ \kappa_{t-1}-\left(k_t-k_{t-1}\right)\pi_t.
\end{aligned}
$$

We therefore have

$$
\begin{aligned}
E\left\{\lambda_t\right\} &\simeq \frac{\theta_w}{\left(1-\theta_w\right)^2}E\left\{\omega_t^2\right\}, \\
E\left\{\Delta_t\right\} &\simeq \theta_p\ E\left\{\Delta_t\right\}+\left(1-\theta_p\right)\tau_1^2 E\left\{\kappa_t\right\}+\frac{\theta_p}{1-\theta_p}E\left\{\pi_t^2\right\}, \\
E\left\{\kappa_t\right\} &\simeq \theta_k\ E\left\{\kappa_t\right\}+\left(1-\theta_k\right)\tau_2^2 E\left\{\Delta_t\right\}+\frac{\theta_k}{1-\theta_k}E\left\{\left(\Delta k_t\right)^2\right\}, \\
E\left\{\psi_t\right\} &\simeq \theta_p\theta_k E\left\{\psi_t\right\}-\tau_2\left(1-\theta_k\right)\left[\theta_p+\tau_1\tau_2\left(1-\theta_p\right)\right]E\left\{\Delta_t\right\} \\
&\qquad -\ \tau_1\left(1-\theta_p\right)\theta_k E\left\{\kappa_t\right\}-E\left\{\left(\Delta k_t\right)\pi_t\right\},
\end{aligned}
$$

or, equivalently,

$$B\begin{bmatrix} E\{\Delta_t\} \\ E\{\kappa_t\} \\ E\{\psi_t\} \end{bmatrix} \simeq C\begin{bmatrix} E\left\{\pi_t^2\right\} \\ E\left\{(\Delta k_t)^2\right\} \\ E\left\{z_t^2\right\} \end{bmatrix},$$

with B

$$\equiv \begin{bmatrix} (1-\theta_p) & -(1-\theta_p)\tau_1^2 & 0 \\ -(1-\theta_k)\tau_2^2 & (1-\theta_k) & 0 \\ \tau_2(1-\theta_k)\left[\theta_p+\tau_1\tau_2(1-\theta_p)\right] & \tau_1(1-\theta_p)\theta_k & (1-\theta_p\theta_k) \end{bmatrix}$$

and $C \equiv \begin{bmatrix} \frac{\theta_p}{1-\theta_p} & 0 & 0 \\ 0 & \frac{\theta_k}{1-\theta_k} & 0 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$,

or, equivalently,

$$\begin{bmatrix} E\{\Delta_t\} \\ E\{\kappa_t\} \\ E\{\psi_t\} \end{bmatrix} \simeq A\begin{bmatrix} E\left\{\pi_t^2\right\} \\ E\left\{(\Delta k_t)^2\right\} \\ E\left\{z_t^2\right\} \end{bmatrix},$$

with $A \equiv B^{-1}C$ and $z_t = \Delta k_t + \pi_t$. We can therefore rewrite our welfare criterion as

$$\widetilde{\mathcal{W}}^{LI} \simeq \Omega_1\ E\left\{y_t^2\right\} + \Omega_2\ E\left\{c_t^2\right\} + \Omega_3\ E\left\{k_t^2\right\} + \Omega_4\ E\left\{n_t^2\right\} + \Omega_5\ E\left\{\omega_t^2\right\} + \Omega_6\ E\left\{\pi_t^2\right\} + \Omega_7\ E\left\{(\Delta k_t)^2\right\} + \Omega_8\ E\left\{z_t^2\right\}, \tag{19}$$

with

$$\Omega_5 \equiv \Omega_\lambda \frac{\theta_w}{(1-\theta_w)^2},\ \Omega_6 \equiv \left[\Omega_\Delta A_{11} + \Omega_\kappa A_{21} + \Omega_\psi A_{31}\right]$$

$$\Omega_7 \equiv \left[\Omega_\Delta A_{12} + \Omega_\kappa A_{22} + \Omega_\psi A_{32}\right],\ \Omega_8 \equiv \Omega_\psi A_{33}.$$

The last equation is (8) in the text.

References

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

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

Dennis, R. 2007. "Optimal Policy in Rational Expectations Models: New Solution Algorithms." *Macroeconomic Dynamics* 11 (1): 31–55.

Edge, R. M. 2003. "A Utility-Based Welfare Criterion in a Model with Endogenous Capital Accumulation." Finance and Economics Discussion Series No. 2003-66, Board of Governors of the Federal Reserve System.

Erceg, C. J., D. W. Henderson, and A. T. Levin. 2000. "Optimal Monetary Policy with Staggered Wage and Price Contracts." *Journal of Monetary Economics* 46 (2): 281–313.

Galí, J. 2015. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework.* 2nd ed. Princeton, NJ: Princeton University Press.

Hicks, J. R., Sir. 1939. *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory.* Oxford: Clarendon Press.

Keynes, J. M. 1936. *The General Theory of Employment, Interest, and Money.* New York: Macmillan.

Khan, A., and J. K. Thomas. 2008. "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics." *Econometrica* 76 (2): 395–436.

Kiyotaki, N., and J. Moore. 1997. "Credit Cycles." *Journal of Political Economy* 105 (2): 211–48.

Kydland, F. E., and E. C. Prescott. 1982. "Time to Build and Aggregate Fluctuations." *Econometrica* 50 (6): 1345–70.

Neiss, K., and E. Nelson. 2003. "The Real-Interest-Rate Gap as an Inflation Indicator." *Macroeconomic Dynamics* 7 (2): 239–62.

Reiter, M., T. Sveen, and L. Weinke. 2013. "Lumpy Investment and the Monetary Transmission Mechanism." *Journal of Monetary Economics* 60 (7): 821–34.

Rotemberg, J., and M. Woodford. 1997. "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy."

In *NBER Macroeconomics Annual 1997*, ed. B. S. Bernanke and J. J. Rotemberg, 297–361. Cambridge, MA: MIT Press.

Schmitt-Grohé, S., and M. Uribe. 2004. "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function." *Journal of Economic Dynamics and Control* 28 (4): 755–75.

———. 2007. "Optimal Simple and Implementable Monetary and Fiscal Rules." *Journal of Monetary Economics* 54 (6): 1702–25.

Sveen, T., and L. Weinke. 2005. "New Perspectives on Capital, Sticky Prices, and the Taylor Principle." *Journal of Economic Theory* 123 (1): 21–39.

———. 2007. "Lumpy Investment, Sticky Prices, and the Monetary Transmission Mechanism." *Journal of Monetary Economics* 54 (Supplement): 23–36.

———. 2009. "Firm-Specific Capital and Welfare." *International Journal of Central Banking* 5 (2): 147–79.

Svensson, L. E. O. 2000. "Open-Economy Inflation Targeting." *Journal of International Economics* 50 (1): 155–83.

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

Thomas, J. K. 2002. "Is Lumpy Investment Relevant for the Business Cycle?" *Journal of Political Economy* 110 (3): 508–34.

Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton, NJ: Princeton University Press.

———. 2005. "Firm-Specific Capital and the New Keynesian Phillips Curve." *International Journal of Central Banking* 1 (2): 1–46.

Disruptions in Large-Value Payment Systems: An Experimental Approach*

Klaus Abbink,[a] Ronald Bosman,[b] Ronald Heijmans,[c]
and Frans van Winden[d]
[a]Monash University
[b]VU University
[c]De Nederlandsche Bank
[d]CREED, University of Amsterdam

This experimental study investigates the individual behavior of banks in a large-value payment system. More specifically, we look at (i) the reactions of banks to disruptions in the payment system, (ii) the way in which the history of disruptions affects the behavior of banks (path dependency), and (iii) the effect of more concentration in the payment system (heterogeneous market versus a homogeneous market). The game used in this experiment is a stylized version of a model of Bech and Garratt (2006) in which each bank can choose between paying in the morning (efficient) or in the afternoon (inefficient). The results show that there is significant path dependency in terms of disruption history. Also, the chance of disruption influences the behavior of the participants. Once the system is moving towards the inefficient equilibrium, it does not easily move back to the efficient one. Furthermore, there is a clear leadership effect in the heterogeneous market.

JEL Codes: C92, D70, D78, E58.

1. Introduction

One of the most significant events in the credit crisis of 2008 was that interbank markets became highly stressed. Liquidity in those markets dried up almost completely because banks suddenly became highly uncertain about each other's creditworthiness. In order to

*Author e-mails: klaus.abbink@monash.edu, r.a.j.bosman@vu.nl, ronald.heijmans@dnb.nl, f.a.a.m.vanwinden@uva.nl.

prevent a collapse of the financial system, central banks intervened by injecting massive volumes of liquidity into the financial system. Our paper relates to stress situations in a particular segment of the financial system, namely large-value payment systems (LVPSs), in which banks pay each other large sums of money during the day.[1] Although during the credit crisis such payment systems were in general functioning properly, disruptions can potentially jeopardize the stability of the financial system as a whole.

The terrorist attacks on the World Trade Center (WTC) in 2001 showed that financial systems are vulnerable to wide-scale disruptions of payment systems. The physical damage to property and communication systems made it difficult or even impossible for some banks to execute payments. The impact of the disruption was not limited to the banks that were directly affected. As a result of fewer incoming payments, other banks became reluctant or in some cases even unable to execute payments themselves (due to insufficient available liquidity). As this could have undermined the stability of the financial system as a whole, the Federal Reserve intervened by providing liquidity through the discount window and open market operations.

Because wide-scale disruptions such as in 2001 do not occur very often, there is not much empirical evidence on how financial institutions behave under extreme stress in payment systems. Research has

[1]Historically, the settlement of interbank payments was done through a netting system in which the payments are settled on a net basis once or several times during the settlement day. With the increase of both the number of transactions and the value of these transactions, the settlement risk increased as well. Banks were increasingly concerned about contagion effects in case of unwinding if one participant would not be able to fulfill its obligation at the end of a netting period. To eliminate this settlement risk, central banks typically developed payment systems in which payments are executed at an individual gross basis, so-called real-time gross settlement (RTGS) systems. Payments are settled irrevocably and with finality. The drawback of RTGS systems is that they require more liquidity because payments usually are not synchronized. To smooth the intraday payment flows, central banks provide intraday credit to their banks. This intraday credit is either collateralized (this holds for most countries including European countries) or priced (United States). An example of a large-value payment system is TARGET2, the euro interbank payment system of the Eurosystem which settled daily in 2008 on average EUR 3.126 billion in value with a volume of 348,000 transactions. Over the years both the value and volume have increased significantly.

therefore focused on simulation techniques. For instance, Soramäki, Bech, and Arnold (2007) and Pröpper, van Lelyveld, and Heijmans (2013) investigated interbank payment systems from a network perspective. To study the impact on the (intraday) dynamics of payment flows, in case a bank changes its own behavior or in case of technical problems, simulations are often used. A comprehensive summary of large-value payment system simulations can be found in Heijmans and Heuver (2012). Many of these simulations focus on the question of what the impact is in terms of liquidity, for the other banks in the payment system, if a single participant delays (intentionally or due to technical problems) payments. This changed behavior can be introduced by modifying historical data in order to reflect the stress scenario under investigation, which requires assumptions on the behavior. Another method that gains momentum in payment systems is agent-based modeling. Our paper is linked to this literature.

Arciero et al. (2009), e.g., present an agent-based model of a real-time gross settlement payment system. Their focus is to provide a methodological contribution, showing how the chosen modeling approach can be employed alongside other tools to study the generation and propagation of systemic risk in payment systems. The model studied by Galbiati and Soramäki (2011) looks at how much liquidity should be posted in a collateralized RTGS system. In contrast to our paper, they do not look at disruptions in payment systems. The approach of our paper is to study disruptions in payment systems in an experimental setting. Basically, we let subjects play a stylized game in a laboratory—which is taken to represent a real LVPS in an essential way—and study interventions in this game. Although subjects in our experiment are students rather than professional managers of LVPSs, there is evidence that students do not behave much different from professionals in economic experiments (e.g., Potters and Van Winden 2000 or Fréchette 2015).[2]

[2]A related issue is that subjects in experiments are paid a relatively modest amount of money compared with professionals in the real world. Note, however, that we paid student subjects on average their opportunity costs of time. In this sense the incentives for the subjects are comparable to those outside the laboratory. Furthermore, there is experimental evidence that increasing the financial stakes does not change behavior much in market and coordination experiments, although in some cases the variance is somewhat reduced (see Jenkins et al. 1998; Kocher, Martinsson, and Visser 2008).

We therefore believe that our results are relevant for understanding real LVPSs. Critics of laboratory experimentation in economics argue that people's behavior in the lab is specific to the experimental situation and may be unconnected to their behavior in the field. They therefore question the external validity of the experimental results. Harrison and List (2004) discuss many of these criticisms and provide a survey of their validity. They plead for conducting experiments both in the lab and in the field. In his response to critics, Camerer (2011) reviews a number of lab experiments in various research areas (among which are experiments on pricing and risk-taking) and concludes that overall they generalize well between lab and field. We have reason to assume therefore that our experiment is relevant for understanding real LVPSs.

An advantage of an experiment is that disruptions can be carefully controlled by the experimenter while the behavioral reactions to these disruptions are determined endogenously (in contrast to simulations where such reactions are assumed). To the best of our knowledge, large-value payment systems have not been studied in the laboratory before, which makes this experiment unique in its kind. McAndrews and Rajan (2000), McAndrews and Potter (2002), and Bech and Garratt (2003) argue that banks' decisions in the U.S. payment system Fedwire can essentially be interpreted as a coordination game. As a vehicle of research we therefore use a stylized game theoretical model developed by Bech and Garratt (2006). In most payment systems participants can execute payments throughout the whole business day. In this model, however, a player has to choose either to pay in the morning, which is considered efficient, or to pay in the afternoon, which is inefficient.[3] This game has two equilibria, one of which is efficient.

These equilibria can be interpreted in a real-life LVPS as follows. Central banks operating an LVPS monitor the functioning of this LVPS closely and encourage banks to pay early (as soon as the obligation is there). One measure central banks have taken to encourage banks to make payments early is to provide collateralized overdrafts. Banks can obtain credit from the central bank without having to pay for this credit. By the end of the day this credit has

[3]This coordination game is known as the stag-hunt game (see also Bech and Garratt 2003).

to be repaid or otherwise the banks do have to pay interest. Furthermore, central banks monitor queues in the LVPS. A payment instruction gets into the queue when the bank has insufficient liquidity available. The payment will be settled when that bank receives sufficient incoming payments or it provides more collateral to be able to receive more (intraday) credit. If a bank systematically delays payments, central banks (may) use moral suasion to "encourage" earlier (timely) payment. In the United Kingdom banks have to live up to so-called throughput guidelines in which banks on average over the month have to settle 50 percent of the value of payments by noon and 75 percent by 2:30 p.m. (see Ball et al. 2011).[4] The reason why central banks monitor their LVPS so closely is that they are afraid that the LVPS will get into a so-called gridlock, i.e., a situation where several payments each await settlement of the others. This might have happened during the attacks on the WTC if the Federal Reserve did not intervene. Given the importance of intraday liquidity, the Bank for International Settlements describes monitoring tools for intraday liquidity management (Basel Committee on Banking Supervision 2013).

The relevance of delaying payments—often called free-riding and in this experiment "paying in the afternoon," is also addressed by Diehl (2013). He looks at measures of individual banks' free-riding behavior in the German part of the European large-value payment system. The difficulty with measuring intentional delay from the payment system data is that only the time the payment was sent to the LVPS and when it was settled is known. This is not necessarily the time the payment had to be paid. However, sudden changes (delays) in payment systems are not likely. Diehl (2013) does not find evidence of free-riding behavior.

Heijmans and Heuver (2014) study the behavior of banks in the payments system before and during the credit crisis which started in the summer of 2007. They find that the timing of payments is an important indicator for identifying potential liquidity problems. They show that some banks indeed delay payments. Besides, they find that some banks limit their exposure to a counterparty by

[4]As described in Becher, Galbiati, and Tudela (2008), enforcement of the throughput guidelines in CHAPS (Clearing House Automated Payment System) currently relies on peer pressure rather than financial or regulatory sanctions.

setting bilateral limits.[5] These limits were set tighter by some of these banks in case there were clear signals about the deteriorating financial soundness of a certain counterparty or even rumors. Setting such limits is the action of a single bank, based on its own risk perception of the market, and not a coordinated action of all the banks simultaneously. This is in contrast to the LIBOR scandal, which could be seen as a more coordinated action of banks. This scandal arose when it was discovered that banks intentionally manipulated their rates either to profit from trades or to disguise their deteriorating creditworthiness. The loans with manipulated rates are usually settled in LVPSs. However, the number of bank transactions related to money market deals is very limited.[6,7] Besides, the timing of the payment of these "manipulated" loans does not have to be affected in any way by this scandal. Therefore, we leave coordinated action out of the scope of this experiment and focus on the individual decisions of banks.[8]

Our study is closely related to the experimental literature on coordination games. Pure coordination games involve multiple equilibria with the same payoff consequences, provided all players choose the same action. The players' task is to take cues from the environment to identify focal points (Schelling 1960; Mehta, Starmer, and Sugden 1994). More akin to our problem are studies on games with Pareto-ranked equilibria. In these games one equilibrium yields higher payoffs to all players than the other equilibriums yield, such that rational players should select it (Harsanyi and Selten 1988). However, experimental subjects often coordinate on inferior equilibria, in particular when the Pareto-dominant equilibrium is

[5]These limits are the maximum debit position a bank is willing to have to a certain counterparty. Bilateral discussion with commercial banks shows that some banks indeed have sophisticated systems available to set and change these bilateral limits. Banks have the opportunity to set such limits in the European large-value payment system, TARGET2. However, most banks choose to set these limits in their internal systems.

[6]Armantier and Copeland (2012) and Arciero et al. (2016) shed light on the euro area and the American unsecured money market during the crisis, respectively.

[7]Besides money market deals, banks settle many other obligations, such as those on behalf of their own business (not money market related) and their clients.

[8]We have excluded communication between players as a first approach. For our study it is important that banks do not know whether the delay of another bank is due to a technical disruption or strategic behavior.

risky (Van Huyck, Battalio, and Beil 1990, 1991), as is the case in our vehicle of research, or other equilibria are more salient (Abbink and Brandts 2008). For an overview of coordination game experiments, see Devetag and Ortmann (2007). None of the existing studies tackles the problem of random disruptions.

Our main research question is how behavior in the payment system is affected by different random disruptions. We define a disruption as a situation where one or more players are unable to execute a payment in a timely manner—for example, because of an individual technical failure or (temporary) financial problems. In addition, we investigate whether concentration in the interbank payment system—in the sense that players are heterogeneous in terms of their size and, consequently, in the risk they may impose on other players—matters. From an economic point this is relevant because consolidation in the financial sector has led to the emergence of a few very large financial institutions.[9] Real payment systems are often characterized by a few large banks and many smaller ones, which make them look like a heterogeneous market.[10] However, the core of the payment system, comprising large banks that together often have a market share of more than 75 percent, looks more like a homogeneous market.[11] There are also LVPSs that are highly tiered. Tiered systems are characterized by a relatively small number of banks participating in the LVPS. These banks also settle payments on behalf of many other banks. Such a tiered system looks similar to a homogeneous market.[12] This means that payment systems can have characteristics of both types of markets. Finally, this paper investigates whether there is any path dependency, taking into account the history of disruption.

[9]The credit crisis has even enhanced this consolidation process. In the United States, for example, investment banks have typically merged with commercial banks. In general, there is a tendency for weaker banks to be taken over by stronger (larger) banks.

[10]The large-value payment system of the Eurosystem (TARGET2) and the United States (Fedwire) are good examples of heterogeneous markets.

[11]For example, in the Dutch part of the European large-value payment system TARGET2, which consists of fifty credit institutions, the five largest banks account for 79 percent of the total value of outgoing daily payments. The thirty-eight smallest ones only cover 5 percent of this value.

[12]The large-value payment system of the United Kingdom (CHAPS) is a good example of a highly tiered system.

The organization of this paper is as follows. Section 2 describes the experimental design (including the game theoretical model), the procedures used, and the predictions. Section 3 discusses the results. Section 4 goes into some policy issues and provides a conclusion.

2. Experimental Design and Procedures

2.1 Design

Our design is based on a model by Bech and Garratt (2006), which is an n-player liquidity management game. The game envisions an economy with n identical banks, which use a real-time gross settlement system operated by the central bank to settle payments and securities. Banks intend to minimize settlement cost. In this game the business day consists of two periods in which banks can make payments: morning or afternoon. At the beginning of the day banks have a zero balance on their accounts at the central bank. At the start of each business day each bank has a request from customers to pay a customer of each of the other (n – 1) banks an amount of Q as soon as possible. To simplify the model, the bank either processes all n – 1 payments in the morning or in the afternoon. If a bank does not have sufficient funds to execute a payment, it can obtain intraday credit, which is costly and reflected by a fee F. This fee can be avoided by banks by delaying their payments to the afternoon. With this delay, however, there are some social and private costs involved, indicated by D. For example, a delay may displease customers or counterparties, costing banks in terms of potential claims and reputation risk. Also, in the case of operational disruptions, payments might not be settled by the end of the business day. This disruption can either be a failure at the payment system to operate appropriately or a failure at the bank itself. The costs in this case can, for example, be claims as a result of unsettled obligations or loss of reputation. The trade-off between the cost F of paying in the morning and cost D of paying in the afternoon is made by each bank individually. Bech and Garratt (2006) investigate the strategic adjustment banks make in response to temporary disruptions. In particular, they focus on equilibrium selection after the disruption is over.

In our experiment we use a simple version of the theoretical model by Bech and Garratt. Because $F \geq D$, there are two equilibria in pure strategies, assuming each bank maximizes its own earnings. Either all banks pay in the morning or all banks pay in the afternoon. The morning equilibrium is the efficient equilibrium.[13] In each of the several rounds of the experiment, the banks have to make a choice between paying in the morning (labeled choice X) and paying in the afternoon (labeled choice Y). In each round, furthermore, there is a known probability p that a bank is forced to pay in the afternoon.[14] This means that the bank cannot pay in the morning but is forced to delay payment to the afternoon. The other banks observe that there was a delay at this bank, but they do not know whether it was caused by a disruption (a forced Y) or a deliberate decision. The probability of disruption is the core parameter of the experiment. After each round, all banks see the choice of the other banks. However, it is not known by the other banks whether a bank was forced to pay in the afternoon or chose to do so intentionally.

The experiment consists of three parts, each consisting of thirty rounds.[15] The probability p varies between the three parts. Instructions for each part were only provided when the respective part began. From a game-theoretic point of view, our design is a finitely repeated version of the (deterministic) game considered by Bech and Garratt. In our setup this should not change the equilibria. Repetition causes additional room for equilibria in games that have Pareto-improving non-equilibrium outcomes in the stage game (Benoit and Krishna 1985). Players can then use equilibrium selection towards the end as a means to enforce more profitable play in the supergame. However, this is not the case in our experiment, as there is no

[13]See proposition 1 of Bech and Garratt (2006).

[14]The inability of many banks during the attacks at the WTC can be interpreted as the forced delayed payment.

[15]A round in the experiment can be interpreted as a business day. Having thirty rounds in one block is a period of constant probability that banks cannot pay early. A (steep) increase of this probability could be seen as a period, e.g., during the attacks at the WTC, in which it took some banks several days to solve the technical difficulties they faced. In a pilot we also investigated a disruption probability of 0 percent. This leads to X choices only. An increase to a 15 percent probability of disruption in the second block of thirty rounds also leads to X choices only—provided that a player has a choice. The pilot showed that there will be consistent coordination on X when the disruption probability is 0 percent.

Figure 1. Two Types of Markets

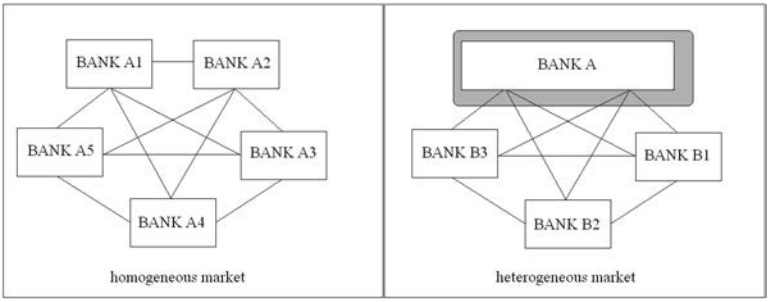


Table 1. Overview of Experimental Treatments

Treatment Name	Type of Market	Disruption Probability p			Number of Groups
		Part 1	Part 2	Part 3	
HOM_15-30-15	Homogeneous	15%	30%	15%	16
HOM_30-15-30	Homogeneous	30%	15%	30%	16
HOM_15-45-15	Homogeneous	15%	45%	15%	15
HOM_45-15-45	Homogeneous	45%	15%	45%	15
HET_15-30-15	Heterogeneous	15%	30%	15%	17
HET_30-15-30	Heterogeneous	30%	15%	30%	14

strategy combination that could improve upon the outcome in the efficient equilibrium.

The experiment investigates the impact of the disruption probability in two types of markets: a homogeneous market and a heterogeneous market. The homogeneous market represents a market in which all banks are identical both in size and impact (n = 5). The heterogeneous market case, on the other hand, constitutes a market in which one bank is twice as large as the other banks, thus making and receiving twice as many payments (n = 4). Conceptually, one can see the heterogeneous market as the homogeneous market where two identical (small) banks have merged; see figure 1. Table 1 provides an overview of the different treatments investigated in the experiment. Instructions and computer screenshots are presented in the appendix.

Table 2. Earnings Table of Homogeneous Market (in experimental currency)

Number of Other Players Choosing X	Number of Other Players Choosing Y	Your Earnings from Choosing X	Your Earnings from Choosing Y
4	0	5	2
3	1	3	2
2	2	1	2
1	3	−1	2
0	4	−3	2

Table 2 shows the earnings in the case of a homogeneous market with five identical banks, where X stands for paying in the morning and Y for paying in the afternoon. Earnings are determined by a fixed payoff of 5, while F = 2 and D = 3/4.[16]

2.2 Procedures

The experiment was run with undergraduate students of the University of Amsterdam using the CREED laboratory. Upon arrival, participants were randomly seated in the laboratory. Subsequently, the instructions for the experiment were given. Students could only participate in the experiment once.

The computerized experiment was set up in an abstract way, avoiding suggestive terms like banks. Choices were simply labeled X and Y. Forced choices were indicated by Y_f on the computer screen of participants. Participants were randomly divided into groups whose composition did not change during the experiment. Participants were labeled A1 to A5 in the homogeneous market and A, B1, B2,

[16]In the case of paying in the afternoon, earnings equal $-(n - 1) \cdot D + 5$, with n being the total number of banks. Earnings if the bank instead chooses paying in the morning equal $-(n - 1 - |Si|m) \cdot F + 5$, where $|Si|m$ denotes the number of other banks paying in the morning. The heterogeneous market case follows straightforwardly and is therefore left out, to save space (see the instructions in the appendix for details). We only notice here that, due to the concentration, the banks run a differential risk in the heterogeneous case. For example, for a small bank only two Y choosers (instead of three) can make the choice of Y optimal now.

and B3 in the heterogeneous market.[17] Note that in the latter market A refers to the large bank (see figure 1). Whether a participant represented a large or a small bank was determined randomly. All payoffs were in experimental talers, which at the end of the experiment were converted into euros at a fixed exchange rate known to the participants. Each experiment took approximately one hour and the average earnings were EUR 18.82, including a show-up fee of EUR 5. In total, 434 students participated in the experiment.

2.3 Predictions

The experimental game has two equilibria in pure strategies when the probability of disruption is "low" (15 percent) or "intermediate" (30 percent). In the first equilibrium, all banks pay in the morning. In the second equilibrium, all banks defer their payment to the afternoon. Note that the first equilibrium is efficient. In this equilibrium all banks are better off than in the second equilibrium. So, one would expect that banks would try to coordinate on this equilibrium. The efficient equilibrium, however, is risky in the sense that paying in the morning is costly when two or more banks decide to defer their payment to the afternoon. Whether or not banks will coordinate on the efficient equilibrium depends, among other things, on their risk attitude. Experimental research shows that in coordination games where the efficient equilibrium is risk-dominated by other equilibria, the efficient equilibrium need not be the obvious outcome (e.g., van Huyck, Battalio, and Beil 1990). When the chance of disruption is "high" (45 percent), there is only one equilibrium, where all banks pay late. In this situation the obvious prediction is that banks coordinate on this equilibrium.[18]

[17] We apply the "unitary player assumption" here, which is common in theoretical and experimental economics. The experimental evidence of a behavioral equivalence of group behavior and individual behavior is very mixed, also regarding markets (see, e.g., Cox and Hayne 2006; Bornstein et al. 2008; Raab and Schipper 2009; Ambrus, Greiner, and Pathak 2013). Bosman, Hennig-Schmidt, and van Winden (2006) discuss eight factors that may account for the mixed evidence. As a first approach, it seems justified therefore to make this assumption.

[18] The finitely repeated version of our (Bech and Garratt) game makes multiple equilibria possible because of the presence of two equilibria—an efficient one (all pay in the morning) and an inefficient one (all pay in the afternoon); see Benoit

The homogeneous and heterogeneous markets in fact have the same two equilibria. From a standard game theoretical point of view, we would expect the same outcome in both markets. From a behavioral point of view it is possible that the outcomes differ. In the heterogeneous market, for example, the large bank may have a disproportionate influence on the behavior of others. Whether such an influence is helpful or harmful in terms of coordinating on the efficient equilibrium is difficult to say a priori, and the experiment will shed more light on such behavioral issues. Finally, we investigate whether there is any path dependency and how this relates to the probability of the disruptions.

3. Results

This section describes the results of the different experimental treatments. We look at plain-choice frequencies and a measure that captures the degree of coordination, called "full coordination" (the situation where participants make the same choice, given that a participant is not forced to "choose" Y). Section 3.1 describes the results for the homogeneous market and section 3.2 describes those for the heterogeneous market.

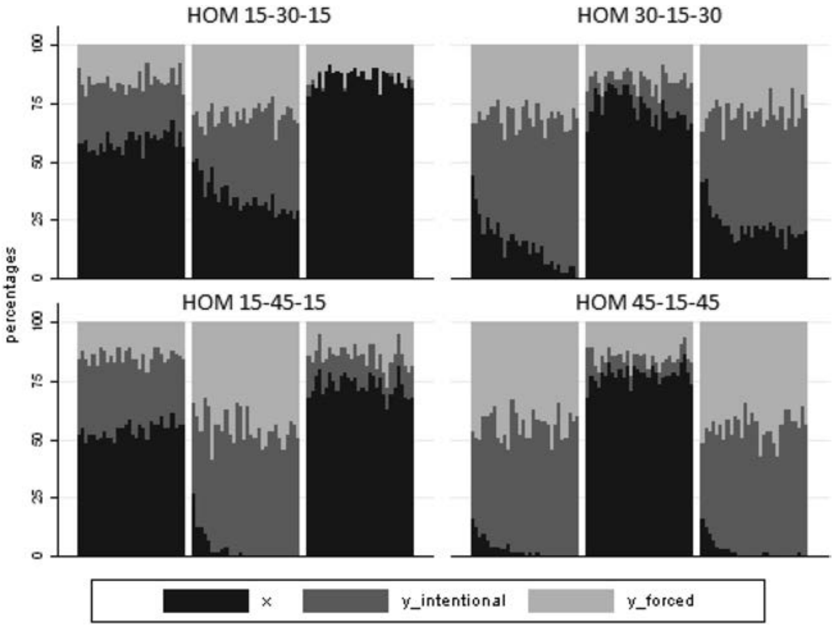
3.1 Homogeneous Market

3.1.1 Choice Frequencies

We take a first shot at the data by simply looking at the choice frequencies of the four homogeneous market treatments, as depicted in figure 1. In figure 2, the HOM_15-30-15 treatment (top left) shows that the choice frequency of X in parts 1 and 3, both with 15 percent disruption probability, does not change much throughout each block of rounds. However, the choice frequency of X in block 3 is higher than in block 1. In fact, intentionally chosen Y in block 3 almost vanishes. These observations suggest that participants learn

and Krishna (1985). Any string of efficient/morning and inefficient/afternoon outcomes can be sustained as equilibrium. However, players cannot reach any higher payoff than that obtained in the efficient (morning) outcome, which makes these additional equilibria not particularly interesting, in contrast to what will happen in the experiment.

Figure 2. Choice Frequencies (homogeneous markets)



Note: The panels of each graph show the choice frequencies per round for the respective parts of the experimental treatment.

to coordinate on the efficient equilibrium over time. Block 2, with a disruption probability of 30 percent, shows that the choice frequency of X decreases from 50 percent to slightly above 25 percent and the choice frequency of intentional Y increases. The results for the reversed-order treatment, HOM_30-15-30 (top right), show a similar pattern for the 30 percent blocks, but a stronger decrease of choice frequency X within the blocks is observed—making the overall choice frequencies of X when p = 30% lower in the reversed-order treatment. This observation suggests that behavior is not fully independent from past disruption experience.

The bottom two graphs, referring to treatments HOM_15-45-15 and HOM_45-15-45, show that a disruption probability of 45 percent quickly leads to choices Y or Y-forced, as predicted. From this it can be concluded that when the disruption probability becomes too large, there is no incentive to choose X anymore, because this

Table 3. Fraction of Groups Fully Coordinating on X or Y (homogeneous markets)

Treatment	Coordination on X			Coordination on Y		
	Part 1	**Part 2**	**Part 3**	**Part 1**	**Part 2**	**Part 3**
HOM_15-30-15	0.56	0.40	0.97	0.19	0.42	0.00
	(0.14)	(0.05)	(0.06)	(0.08)	(0.15)	(0.00)
HOM_30-15-30	0.11	0.76	0.24	0.66	0.08	0.60
	(0.08)	(0.12)	(0.05)	(0.23)	(0.08)	(0.14)
HOM_15-45-15	0.53	0.01	0.80	0.30	0.91	0.11
	(0.14)	(0.04)	(0.06)	(0.08)	(0.14)	(0.04)
HOM_45-15-45	0.01	0.86	0.01	0.86	0.06	0.91
	(0.03)	(0.12)	(0.03)	(0.16)	(0.02)	(0.13)

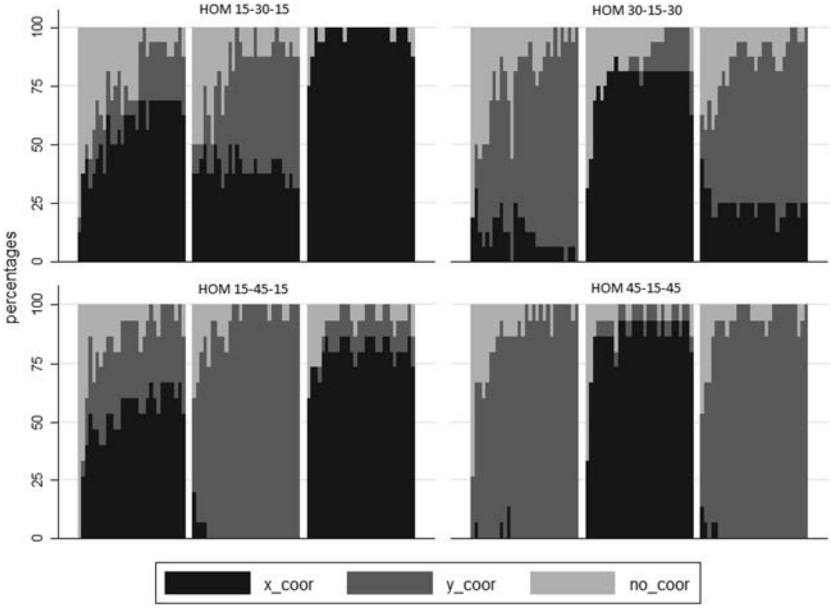
Note: Standard deviation is shown in parentheses.

will lead to losses for the participants. Comparing the bottom left graph with the top left graph shows that the increasing trend in X choices in going from block 1 to 3 is similar. However, in block 3 of HOM_15-45-15, the increase in X appears less strong than in block 3 of HOM_15-30-15.

3.1.2 Frequency of Full Coordination

Table 3 shows the average fraction of the groups that fully coordinate on X and Y for the four homogeneous market treatments. There is full coordination on X or Y when all of the participants within one group who have a choice (i.e., who are not forced to choose Y) choose X or Y, respectively. There has to be at least one participant who has a choice in order to get full coordination on X or Y. Figure 3 shows the level of coordination on X (black bar) and Y (dark grey bar), and the absence of coordination (light grey bar), for each round of the four treatments. The data show that there is more coordination on X when the disruption probability is low ($p = 15\%$) and more coordination on Y when the disruption probability is intermediate or high (30 percent or 45 percent, respectively) ($p < 0.01$, binomial test for block 1 between treatments). In the context of a payment system, this suggests that larger disruptions are associated with less efficiency.

Figure 3. Percentage of Groups Fully Coordinating on X or Y (homogeneous markets)



Note: The panels of each graph show the full coordination per round for the respective parts of the experimental treatment.

Result 1. *A higher disruption probability leads to less coordination on X and more coordination on Y.*

Both table 3 and figure 3 show that there is more coordination either on X or Y in block 3 compared with block 1. There is significantly more coordination on X in the third block compared with block 1 for the HOM_15-30-15, the HOM_30-15-30, and the HOM_15-45-15 treatments and less coordination on Y (all $p < 0.01$, binomial test). Participants thus learn to coordinate on the efficient equilibrium, which is even speeded up if there is a prior disruption chance of 30 percent or 45 percent. The table and figure also show that for a disruption probability of 45 percent, coordination on X almost vanishes and quickly moves to the inefficient equilibrium. Coordination on X only occurs occasionally in the first few rounds. This is in line with the low-choice frequencies of X in that case, presented in the previous subsection. In the context of a payment system, this means that if the disruption is very likely, there is no incentive anymore to pay as soon as possible. This situation seems similar to the one of the attacks on the WTC in 2001 when many banks, including some large ones, were not able to execute payments due to technical problems. Some banks were reluctant to execute any payment, even though they were able to, because they did not know the impact of the attacks on the stability of the financial system. Understandably, these events threatened to move the payments system to the inefficient equilibrium, which was a reason for the authorities to intervene.

Figure 3 further shows that the developments within blocks are somewhat monotonic, especially the blocks with 15 percent disruption probability. With respect to disruption probability of 30 percent and 45 percent, there is a (bit of a) decreasing trend for X. This suggests that once a trend has been established in the payments system, it is unlikely to reverse. The 15 percent disruption probability of HOM_30-15-30 shows a higher level of coordination on X than block 1 of HOM_15-30-15 but a lower level when compared with block 3. Comparing HOM_15-30-15 with HOM_15-45-15 shows that there is no significant difference in coordination on X in block 1. Block 3 of these two treatments, however, shows some differences, with significantly more coordination on X in HOM_15-30-15 ($p < 0.01$, binomial test). Although the disruption probability is the same, the history of disruption exposure differs between these two treatments. The previous block has either a probability of disruption of 30 percent or 45 percent, leading to different behavior. Block 2 of HOM_15-45-15 shows 91 percent coordination on Y and almost 0 percent coordination on X. For HOM_15-30-15 this is 42 percent coordination on Y and 40 percent on X. This suggests that the disruption history is important for the coordination on both X and Y. In terms of payment systems, this means that the payment behavior of banks depends on history.

RESULT 2. *Overall, there is more coordination in the third part of the experiment than in the first part, given the same disruption probability. If the chance of disruption is low ($p = 0.15$) or intermediate ($p = 0.3$) in the first part, there is more coordination on X in the* *third part. If the disruption probability is high (p = 0.45), there is strong coordination on the inefficient equilibrium.*

Result 3. *There is evidence of path dependency, as the outcome depends on the disruption history.*

Confidence between banks is not a static fact, as became clear during the current financial crisis. Banks became reluctant to execute their payments to financial institutions that were "negative in the news." Especially the bankruptcy of Lehman Brothers in October 2008 caused a shock wave of uncertainty through the whole financial system. Banks became aware of the fact that even large (systemically important) banks might not stay in business. The interbank money market, which gives banks with a surplus of liquidity the opportunity to lend money to banks with a temporary shortage, came to a standstill (see, e.g., Armantier and Copeland 2012 and Arciero et al. 2016 for the unsecured money market in the euro area and the United States, respectively). This indicates that recent history is important for the level of confidence banks have in each other, as our experimental result suggests.

3.2 Heterogeneous Market

Recall that in the heterogeneous markets the number of banks is four instead of five. One of the banks is now twice as large in size and impact compared with the other three banks.

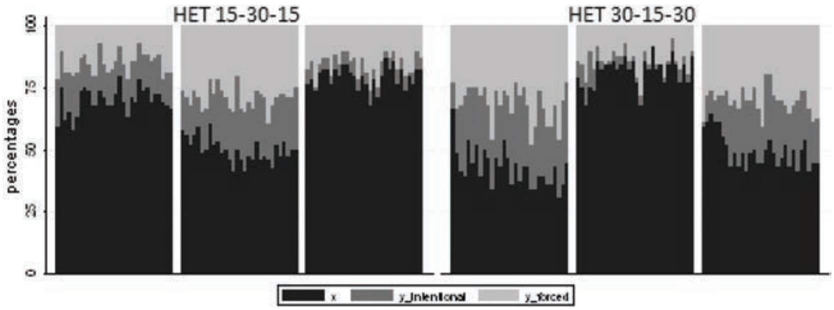
3.2.1 Choice Frequencies

Again we take a first shot at the data by looking at plain-choice frequencies in the two heterogeneous markets (see figure 4 and table 3). The left graph of the figure, concerning treatment HET_15-30-15, shows similar trends as in HOM_15-30-15 (figure 2). However, across the different blocks, participants in the heterogeneous markets show a tendency to choose X more often.

3.2.2 Frequency of Full Coordination

Table 4 shows the average fraction of the groups that fully coordinate on X or Y in the two heterogeneous market treatments, while

Figure 4. Choice Frequencies (heterogeneous markets)



Note: The panels of each graph show the choice frequencies per round for the respective parts of the experimental treatment.

Table 4. Fraction of Groups Fully Coordinating on X or Y (heterogeneous markets)

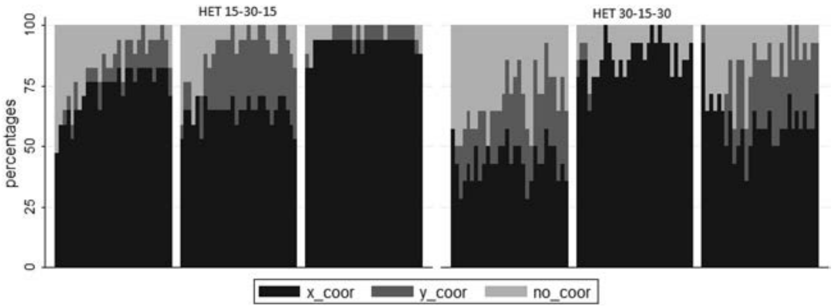
Treatment	Coordination on X			Coordination on Y		
	Part 1	Part 2	Part 3	Part 1	Part 2	Part 3
HET_15-30-15	0.73	0.64	0.93	0.09	0.24	0.04
	(0.10)	(0.05)	(0.03)	(0.05)	(0.09)	(0.03)
HET_30-15-30	0.44	0.87	0.60	0.23	0.01	0.22
	(0.08)	(0.08)	(0.10)	(0.10)	(0.02)	(0.12)

Note: Standard deviation is shown in parentheses.

figure 5 shows the coordination over the rounds. Comparing the full coordination on both X and Y of the heterogeneous market with the corresponding homogeneous one of section 3.1.2 shows that trends between blocks are similar. However, given the same immediate disruption history, there is significantly more coordination on X in the heterogeneous market treatments than in the homogeneous market in five out of the seven possible cases with the same immediate disruption history (all five cases $p < 0.01$, binomial test).[19] In the two

[19] All blocks of treatment 1 and 2 can be compared with treatment 5 and 6, respectively. The first block of treatment 3 can be compared with the first block of treatment 5. Two cases are not significant. These relate to blocks 2 and 3, given an immediate disruption history of 15 percent ($p = 0.2$ and $p = 0.6$, respectively).

Figure 5. Percentage of Groups Fully Coordinating on X or Y (heterogeneous markets)



Note: The panels of each graph show the full coordination per round for the respective parts of the experimental treatment.

other cases, there is no significant difference. Note that only blocks that have the same disruption history are compared. These results suggest that coordination is more prominent in a heterogeneous market with asymmetry between participants. A potential explanation is that there is a leadership effect of the large bank, which may feel more responsible than the small banks to choose X because of its relatively large effect on the earnings of all participants. In terms of payment systems, this suggests that a system that consists of one (or perhaps a few) large banks and many small(er) banks will lead to more efficiency than a payment system in which banks are more similar in size.

Result 4. *The heterogeneous market leads to more coordination on the efficient equilibrium in most situations characterized by the same immediate disruption history.*

To shed more light on this explanation, we look in more detail at whether the small banks follow the large bank or the other way around in both the 15 percent and 30 percent disruption probability cases. Table 5 shows the reaction of the small banks to the choice of the large bank in previous round(s). The table shows that if the large bank chose X in one or more consecutive rounds, counting from the previous round backwards, there is roughly a 90 percent chance that small banks with a choice (no forced Y) will choose X as well.

Table 5. Leadership of Large Bank—Percentage of Small Banks Following the Large Bank

Choice of Large Bank	Choice of Small Banks = X if Choice of Large Banks Is X	Events	Choice of Small Banks = X if Choice of Large Bank Is Y	Events
Only Once in a Row so Far	89%	1,560	83%	1,544
Only Twice in a Row so Far	92%	1,056	63%	516
Only Three Times in a Row so Far	94%	808	39%	216
Only Four Times in a Row so Far	92%	592	20%	144

Note: "So far" means from the previous round counted backwards.

If the large bank has chosen Y, either intentionally or forcedly, the small banks seem to ignore this when it is only once, as they still choose X 83 percent of the time in that case. Possibly, the small banks will reason that the large bank might have been forced and most likely will choose X in the next round again. The number of small banks choosing X quickly drops if the large bank chooses Y more than once in a row. This can be explained by the fact that two or more forced Ys are not very likely and may be a signal of bad intention rather than bad luck. Note from the payoff table in the appendix that in this situation the payoff for the small banks from choosing X becomes markedly lower, in particular when one other small bank also chooses Y. Large banks in an LVPS do generally have more liquidity available at their account than smaller ones, as a result of the reserve maintenance requirements. This liquidity can intraday be used to make payments. In the case of the European LVPS, TARGET2, many banks have sent in payment instructions before the opening of the system (7:00 a.m.). Banks with sufficient liquidity available will settle all payments immediately, while the ones not having sufficient liquidity will have to wait until sufficient liquidity has been received. However, if these large banks are not able to pay or if they delay payments, this may hamper the smooth functioning of the payment system.

The fact that small banks follow the larger bank is consistent with actual behavior in payment systems, where small banks typically depend on the liquidity of the large bank. For example, it is observed in the Netherlands that large banks have a tendency to start paying large amounts right after opening of the payment system, which corresponds to paying in the morning in terms of our experimental game. The smaller banks usually follow immediately after that. This can still be considered as "paying in the morning" because these payments follow almost instantaneously after the payments of the large banks. This means that the large banks provide liquidity to the small ones, which the latter can use to fulfill their payment obligations.

3.3 Dynamics

Our results show that when the probability of disruptions is moderate, subjects typically achieve a high level of coordination on the efficient equilibrium, while with higher probabilities results are more mixed. We have studied several behavioral rules (heuristics) to explain the observed dynamic behavioral patterns. Most prominently, we investigated the "myopic best response" rule which is an application of the adjustment process that Bech and Garratt propose (section 3 of their paper). However, in addition, we have looked at simply copying the behavior of whoever was most successful ("imitation") as an even simpler and (in the literature on behavioral adjustment) often-considered heuristic, as well as a more general behavioral rule where participants choose X as long as it is profitable. As none of these heuristics turns out to be very successful in capturing our data, we do not include them in this paper and refer to section 4 of Abbink et al. (2010) for a detailed analysis and graphical representation of the outcomes.

4. Conclusions

In this paper we used a stylized coordination game of Bech and Garratt (2006) to experimentally study bank behavior in a large-value payment system that is hindered by disruptions. We draw the following conclusions.

First, once behavior moves in the direction of coordination on the inefficient equilibrium, it is not likely that behavior moves back to the efficient equilibrium. The reason for this is that one player has to take the lead in going for the efficient equilibrium, but this is costly if other players do not follow suit. In the context of a payment system, these findings suggest that once a trend has been established, it is unlikely to reverse. In a situation where some banks begin to defer their payments, an intervention from the central bank may be highly desired or even necessary. When banks do not have access to sufficient liquidity—i.e., they are forced to go for the inefficient equilibrium—central banks can use their discount window to relieve market stress, which happened during the attacks at the WTC. If some (critical) banks deliberately delay payments without having liquidity problems, the central bank can use its authority to encourage banks to start paying earlier (cf. Chaudhuri, Schotter, and Sopher 2009, who study the role of advice in coordination games). Such moral suasion only works though if the payment system has not been disturbed totally (i.e., coordinated fully on the inefficient equilibrium). So, once coordination failures start emerging, central banks need to react quickly; otherwise, trust between banks might have fully vanished and coordination on the "good" equilibrium becomes highly unlikely. Note that in our experimental study there was no role for the central bank. We believe that extending the game by allowing central bank interventions would be an interesting avenue for future experimental work.

Second, coordination on the efficient equilibrium turns out to be easier in a heterogeneous market where there is a clear leader in terms of size. If such a leader goes for the efficient equilibrium, 90 percent of smaller players who have a choice follow the leader. If the leader is not cooperative for several rounds in a row (forcedly or deliberately), the smaller players rapidly move to such a strategy as well. Given the critical role of the large player for the system as a whole, it is essential from a payment system perspective to minimize the chance that large banks are not able to execute payments due to their own technical problems. It may therefore be desirable to oblige such critical participants to take extra safety measures with regard to their technical infrastructure.

Finally, our experiment shows that small frictions in coordination games can be absorbed easily and need not jeopardize the stability of the efficient equilibrium (cf. the 15 percent disruption cases). However, when friction becomes larger, the system can move quickly to the undesired equilibrium and stay there. In the context of payment systems, this suggests that it is very important to closely monitor the payment flows of (critical) participants in the system. If deviant payment behavior is observed by one or more participants, it is important to find the reason for this behavior. If the cause is a technical problem of one participant, the other participants in the payment system should be informed about the incident. In this way it may be avoided that the other participants falsely conclude that the deviant behavior is a deliberate action, for example, related to liquidity considerations. Such communication is especially important during times of market stress, when false rumors can easily arise. Since we did not study communication in our experiment, it is an open research question whether this could work or not.

Appendix

Instructions for the Homogeneous Market Case

The instructions for the homogeneous market case are shown below. Between the different experimental treatments, only the percentages change. The instructions listed here are for the 15%-30%-15% case.

Instructions

Welcome to this experiment. The experiment consists of three parts in which you will have to make decisions. In each part it is possible to earn money. How much you earn depends on your own decisions and on the decisions of other participants in the experiment. At the end of the experiment a show-up fee of 5 euros plus your total earnings during the experiment will be paid to you in cash. Payments are confidential; we will not inform any of the other participants. In the experiment, all earnings will be expressed in talers, which will be converted to euros according to the exchange rate: 1 taler = 6 Eurocents.

During the experiment you will participate in a group of five players. You will be matched with the same players throughout the experiment. These other players in your group will be labeled P2, P3, P4, and P5.You will not be informed of who the other players are, nor will they be informed of your identity.

Talking or communicating with others during the experiment is not permitted. If you have a question, please raise your hand and we will come to your desk to answer it.

Warning: In this experiment you can avoid making any loss (negative earnings). However, note that if you end up with a loss, it will be charged against your show-up fee.

We start now with the instructions for part 1, which have been distributed also on paper. The instructions for the other two parts will be given when they start.

Instructions: Part 1

This part consists of thirty rounds. In each round you and the other four players in your group will have to choose one of two options: X or Y. Your earnings in a round depend on your choice and on the choices of the other four players, in the following manner:

- If you choose Y, your earnings are 2 talers regardless of the choices of the others.
- If you choose X, your earnings depend on how many of the other players choose Y.

Your exact earnings in talers from choosing X or Y, for a given number of other players choosing Y, are listed in the following table. This earnings table is the same for all players.

Number of Other Players Choosing Y	**Your Earnings from Choosing X**	**Your Earnings from Choosing Y**
0	5	2
1	3	2
2	1	2
3	−1	2
4	−3	2

For example, if two other players choose Y, then your earnings from choosing X will be 1, while your earnings from choosing Y would be 2.

Forced Y. Note, however, that you may not be free to choose your preferred option. In each round, each of you will face a 15 percent chance that you are forced to choose option Y. We will call this a "forced Y."

Whether or not a player is forced to choose Y is randomly determined by the computer for each player separately and independently from the other players. Further, a forced Y does not depend on what happened in previous rounds.

On the computer screen where you make your decision, you will be reminded of this chance of a forced Y, for your convenience. Furthermore, in the table at the bottom of that screen (showing past decisions and earnings), your forced Ys are indicated in the column showing your choices with an "F." Note that you will not be informed of other players' forced Y choices.

You are now kindly requested to do a few exercises on the computer to make you fully familiar with the earnings table. In these exercises you cannot earn any money. Thereafter, we will start with part 1.

Please raise your hand if you have any question. We will then come over to your table to answer your question.

Instructions Part 2

Part 2 is exactly the same as part 1, except for one modification: In each round, each of you will now face a 30 percent chance that you are forced to choose option Y. Are there any questions?

Instructions Part 3

Part 3 is exactly the same as part 2, except for one modification: In each round, each of you will now face a 15 percent chance that you are forced to choose option Y, as in part 1. Are there any questions?

Instructions for the Heterogeneous Case

The instructions for the heterogeneous market case are shown below. Again, between the different experimental treatments only the percentages change. The instructions listed here are for the 15%-30%-15% case.

Instructions

Welcome to this experiment. The experiment consists of three parts in which you will have to make decisions. In each part it is possible to earn money. How much you earn depends on your own decisions and on the decisions of other participants in the experiment. At the end of the experiment a show-up fee of 5 euros plus your total earnings during the experiment will be paid to you in cash. Payments are confidential; we will not inform any of the other participants. In the experiment, all earnings will be expressed in talers, which will be converted to euros according to the exchange rate: 1 taler = 6 Eurocents.

During the experiment you will participate in a group of four players. You will be matched with the same players throughout the experiment. There are two types of players: A and B. The difference is related to the consequences of their decisions, as will be explained below. In fact, there will be one A player and three B players in your group. If you happen to be player A, then the others are B players, who will be labeled B1, B2, and B3. If you are a B player, then the other players in your group comprise a player A and two other B players, denoted as B2 and B3. You will learn your type when Part 1 starts; it will stay the same during the whole experiment. Because we have pre-assigned a type to each table, you have drawn your type yourself when you selected a table number in the reception room. You will not be informed of who the other players are, nor will they be informed of your identity.

Talking or communicating with others during the experiment is not permitted. If you have a question, please raise your hand and we will come to your table to answer it.

We start now with the instructions for part 1, which have been distributed also on paper. The instructions for the other two parts will be given when they start.

Instructions Part 1

First of all, note that your type (A or B) will be shown at the upper-left part of your computer screen, below a window showing the round number.

This part consists of thirty rounds. In each round you and the other three players in your group will have to choose one of two options: X or Y. Your earnings in a round depend on your type (A or B), your choice, and the choices of the other three players, in the following manner:

- If you choose Y, your earnings are 2, regardless of your type and the choices of the others.
- If you choose X, your earnings depend on your type and on how many of the other players choose Y.

Your exact earnings from choosing X or Y, given your type and the Y choices of the other players in your group, are listed in the following tables for, respectively, player A and a B player.

The following are some examples, for illustration.

Suppose you are a player A, and you choose X while one of the other players chooses Y. Then the upper table shows that your earnings will be 3.

Alternatively, suppose you are a B player, and you choose X while one of the other players chooses Y. Then it depends on whether this other player choosing Y is a player A or another B player. If it is player A, then the lower table shows that your earnings are 1, while your earnings are 3 if it is a B player. Thus, player A has a larger impact on your earnings than a B player.

Player A

Your Choice	**Number of B Players Choosing Y**	**Your Earnings**
X	0	5
X	1	3
X	2	1
X	3	−1
Y	0	2
Y	1	2
Y	2	2
Y	3	2

Player B

Player A's Choice	Number of Other B Players Choosing Y	Your Earnings from Choosing X	Your Earnings from Choosing Y
X	0	5	2
X	1	3	2
X	2	1	2
Y	0	1	2
Y	1	−1	2
Y	2	−3	2

Forced Y. Note, however, that you may not be free to choose your preferred option. In each round, each of you will face a 15 percent chance that you are forced to choose option Y. We will call this a "forced Y."

Whether or not a player is forced to choose Y is randomly determined by the computer for each player separately and independently from the other players. Further, a forced Y does not depend on what happened in previous rounds.

On the computer screen where you make your decision, you will be reminded of this chance of a forced Y, for your convenience. Furthermore, in the table at the bottom of that screen (showing past decisions and earnings), your forced Ys are indicated in the column showing your choices with an "F." Note that you will not be informed of other players' forced Y choices.

You are now kindly requested to do a few exercises on the computer to make you fully familiar with the earnings table. In these exercises you cannot earn any money. Thereafter, we will start with part 1.

Please raise your hand if you have any question. We will then come over to your table to answer your question.

Instructions Part 2

Part 2 is exactly the same as part 1, except for one modification: In each round, each of you will now face a 30 percent chance that you are forced to choose option Y. Are there any questions?

Instructions Part 3

Part 3 is exactly the same as part 2, except for one modification: In each round, each of you will now face a 15 percent chance that you are forced to choose option Y, as in part 1. Are there any questions?

References

Abbink, K., R. Bosman, R. Heijmans, and F. Van Winden. 2010. "Disruptions in Large Value Payment Systems: An Experimental Approach." DNB Working Paper No. 263.

Abbink, K., and J. Brandts. 2008. "Pricing in Bertrand Competition with Increasing Marginal Costs." *Games and Economic Behavior* 63 (1): 1–31.

Ambrus, A., B. Greiner, and P. Pathak. 2013. "How Individual Preferences Get Aggregated in Groups — An Experimental Study." Working Paper.

Arciero, L., C. Biancotti, L. D'Aurizo, and C. Impenna. 2009. "Exploring Agent-Based Methods for the Analysis of Payment Systems: A Crisis Model for StarLogo TNG." *Journal of Artificial Societies and Social Simulations* 12 (1).

Arciero, L., R. Heijmans, R. Heuver, M. Massarenti, C. Picillo, and F. Vacirca. 2016. "How to Measure the Unsecured Money Market: The Eurosystem's Implementation and Validation using TARGET2 Data." *International Journal of Central Banking* 12 (1): 247–80.

Armantier, O., and A. M. Copeland. 2012. "Assessing the Quality of 'Furfine-Based' Algorithms." Staff Report No. 575, Federal Reserve Bank of New York.

Ball, A., D. Denbee, M. Manning, and A. Wetherilt. 2011. "Intraday Liquidity: Risk and Regulation." Financial Stability Paper No. 11, Bank of England.

Basel Committee on Banking Supervision. 2013. "Monitoring Tools for Intraday Liquidity Management." Bank for International Settlements (April).

Bech, M., and R. Garratt. 2003. "The Intraday Liquidity Management Game." *Journal of Economic Theory* 109 (2): 198–210.

———. 2006. "Illiquidity in the Interbank Payment System Following Wide-Scale Disruptions." Staff Report No. 239, Federal Reserve Bank of New York.

Becher, C., M. Galbiati, and M. Tudela. 2008. "The Timing and Funding of CHAPS Sterling Payments." *Economic Policy Review* (Federal Reserve Bank of New York) 14 (2, September): 113–33.

Benoit, J., and V. Krishna. 1985. "Finitely Repeated Games." *Econometrica* 53 (4): 905–22.

Bornstein, G., T. Krugler, D. Budescu, and R. Selton. 2008. "Repeated Price Competition between Individuals and between Teams." *Journal of Economic Behavior and Organization* 66 (3–4): 808–21.

Bosman, R., H. Hennig-Schmidt, and F. van Winden. 2006. "Exploring Group Decision Making in a Power-to-Take Experiment." *Experimental Economics* 9 (1): 35–51.

Camerer, F. 2011. "The Promise and Success of Lab-Field Generalizability in Experimental Economics: A Critical Reply to Levitt and List." In *Handbook of Experimental Economic Methodology*, ed. G. R. Fréchette and A. Schotter (chapter 14). Oxford University Press.

Chaudhuri, A., A. Schotter, and B. Sopher. 2009. "Talking Ourselves to Efficiency: Coordination in Inter-generational Minimum Effort Games with Private, Almost Common and Common Knowledge of Advice." *Economic Journal* 119 (534): 91–122.

Cox, J., and S. Hayne. 2006. "Barking up the Right Tree: Are Small Groups Rational Agents?" *Experimental Economics* 9 (3): 209–22.

Devetag, G., and A. Ortmann. 2007. "When and Why? A Critical Survey on Coordination Failure in the Laboratory." *Experimental Economics* 10 (3): 311–44.

Diehl, M. 2013. "Measuring Free Riding in Large-Value Payment Systems: The Case of TARGET2." *Journal of Financial Market Infrastructures* 1 (3): 31–53.

Fréchette, G. 2015. "Laboratory Experiments: Professionals versus Students." In *Handbook of Experimental Economic Methodology*, ed. G. R. Fréchette and A. Schotter, 36–90 (chapter 17). Oxford University Press.

Galbiati, M., and K. Soramäki. 2011. "An Agent-based Model of Payment Systems." *Journal of Economic Dynamics and Control* 35 (6): 859–75.

Harrison, G., and J. List. 2004. "Field Experiments." *Journal of Economic Literature* 42 (4): 1009–55.

Harsanyi, J., and R. Selten. 1988. *A General Theory of Equilibrium Selection in Games.* MIT Press.

Heijmans, R., and R. Heuver. 2012. "Preparing Simulations in Large Value Payment Systems using Historical Data." In *Simulation in Computational Finance and Economics: Tools and Emerging Applications*, ed. B. Alexandrove-Kabadjova, S. Martinez-Jaramillo, A. Garcia-Almanza, and E. Tsang, 46–68 (chapter 3). IGI Global.

———. 2014. "Is this Bank Ill? The Diagnosis of Doctor TARGET2." *Journal of Financial Market Infrastructures* 2 (3): 31–53.

Jenkins, G. D., Jr., A. Mitra, N. Gupta, and J. D. Shaw. 1998. "Are Financial Incentives Related to Performance? A Meta-Analytic Review of Empirical Research." *Journal of Applied Psychology* 83 (5): 777–87.

Kocher, M., P. Martinsson, and M. Visser. 2008. "Does Stake Size Matter for Cooperation and Punishment?" *Economics Letters* 99 (3): 508–11.

McAndrews, J., and S. Potter. 2002. "Liquidity Effects of the Events of September 11, 2001." *Economic Policy Review* (Federal Reserve Bank of New York) 8 (2, November): 59–79.

McAndrews, J., and S. Rajan. 2000. "The Timing and Funding of Fedwire Funds Transfers." *Economic Policy Review* (Federal Reserve Bank of New York) 6 (2, July): 17–32.

Mehta, J., C. Starmer, and R. Sugden. 1994. "The Nature of Salience: An Experimental Investigation of Pure Coordination Games." *American Economic Review* 84 (3): 658–73.

Potters, J., and F. van Winden. 2000. "Professionals and Students in a Lobbying Experiment: Professional Rules of Conduct and Subject Surrogacy." *Journal of Economic Behavior and Organization* 43 (4): 499–522.

Pröpper, M., I. van Lelyveld, and R. Heijmans. 2013. "Network Dynamics of TOP Payments." *Journal of Financial Market Infrastructures* 1 (3): 3–29.

Raab, P., and B. Schipper. 2009. "Cournot Competition between Teams: An Experimental Study." *Journal of Economic Behavior and Organization* 72 (2): 691–702.

Schelling, T. C. 1960. *The Strategy of Conflict*. Harvard University Press.

Soramäki, K., M. Bech, and J. Arnold. 2007. "The Topology of Interbank Payment Flows." *Physica A* 379 (1): 317–33.

Van Huyck, J., R. Battalio, and R. Beil. 1990. "Tacit Coordination Games, Strategic Uncertainty and Coordination Failure." *American Economic Review* 80 (1): 234–48.

———. 1991. "Strategic Uncertainty, Equilibrium Selection, and Coordination Failure in Average Opinion Games." *Quarterly Journal of Economics* 106 (3): 885–910.

Cyclicality and Firm Size in Private Firm Defaults*

Thais Lærkholm Jensen,[a] David Lando,[b] and Mamdouh Medhat[c]

[a]University of Copenhagen and Danmarks Nationalbank

[b]Copenhagen Business School

[c]Cass Business School – City, University of London

The Basel II/III and CRD IV Accords reduce capital charges on bank loans to smaller firms by assuming that the default probabilities of smaller firms are less sensitive to macroeconomic cycles. We test this assumption in a default intensity framework using a large sample of bank loans to private Danish firms. We find that controlling only for size, the default probabilities of small firms are, in fact, less cyclical than the default probabilities of large firms. However, accounting for firm characteristics other than size, we find that the default probabilities of small firms are equally cyclical or even more cyclical than the default probabilities of large firms. These results hold using a multiplicative Cox model as well as an additive Aalen model with time-varying coefficients.

JEL Codes: G21, G28, G33, C41.

1. Introduction

Small and medium-sized enterprises (SMEs) typically depend more heavily on funding from banks than do larger firms. It is therefore conceivable that SMEs are hit harder during a financial crisis in

*We are grateful to a Danish financial institution for providing data on loans to private firms. We thank Diana Bonfim, Laurence Deborgies-Sanches (discussant), Darrell Duffie, Johan Hombert, Harrison Hong (editor), Allan Mortensen, Lasse Heje Pedersen, Hanne Schultz, and attendees at FEBS/LabEx-ReFi 2013 for helpful discussions and suggestions. Support from the Center for Financial Frictions (FRIC), grant no. DNRF102, and Danmarks Nationalbank, is gratefully acknowledged. Corresponding author (Lando): E-mail: dl.fi@cbs.dk. Tel.: +45 3815 3613. Address: Solbjerg Plads 3, DK-2000, Frederiksberg, Denmark.

which banks' capital constraints are binding. As a way to facilitate bank funding of SMEs, the Basel II Accord prescribes lower capital charges for loans to the SME segment. Technically, the reduction in capital charges is implemented by prescribing a lower asset correlation to be used when calculating capital charges. To the extent that asset correlation arises because of common dependence on macroeconomic shocks, the reduction corresponds to assuming that the default probabilities of SMEs are less sensitive to macroeconomic cycles than those of larger firms. These reductions in capital charges were recently reaffirmed and extended in the Basel III Accord and in the fourth Capital Requirements Directive (CRD IV).[1]

This paper uses a large sample of loans to private Danish firms to test whether there is empirical support for the assumption that the default probabilities of smaller firms are less cyclical. Our results indicate that solely discriminating with respect to firm size, the default probabilities of small firms are in fact less sensitive to macroeconomic cycles compared with the default probabilities of large firms. However, when we account for differences in firm characteristics other than size, our results indicate that the default probabilities of small firms are as cyclical or even more cyclical than the default probabilities of large firms. This indicates that the size effect arises because of omitted variables. These results are robust to different regression models and different ways of dividing our sample into small and large firms. The evidence from our data suggests that a bank that properly accounts for firm characteristics in its default risk modeling will not experience a weaker effect of economic cycles in the SME segment, and hence the reduced capital charge will in fact make a bank with a high exposure to the SME segment more risky.

According to the European Commission's (2016) annual report, SMEs comprised over 99 percent of all enterprises, accounted for 67 percent of total employment, and stood for over 70 percent of

[1] Article 273 of the Basel Committee on Banking Supervision's (2006) report states the following: "Under the IRB approach for corporate credits, banks will be permitted to separately distinguish exposures to SME borrowers (defined as corporate exposures where the reported sales for the consolidated group of which the firm is a part is less than €50 million) from those to large firms. A firm-size adjustment ... is made to the corporate risk weight formula for exposures to SME borrowers."

job growth in the non-financial business sector in the twenty-eight European Union countries. Hence their importance for the economy provides a political justification for the separate treatment. Furthermore, shifting banks' risk exposure towards smaller firms could be in the interest of regulators—for instance, because larger firms are more likely to have larger loans or loans across several banks, which could create contagion within and across banks in case of defaults. But such effects are already accounted for using so-called granularity adjustments. Our focus is on whether an additional adjustment of capital charges through lower correlation with the macroeconomic environment can be justified empirically.

Our data cover the period 2003–12 and consist of obligor, loan, and default information for a large sample of private Danish firms. We devise two different ways of testing whether defaults of smaller firms are more or less cyclical. First, we use the standard multiplicative Cox regression model to estimate the effects of firm-specific and macroeconomic variables on default intensities. Our Cox models confirm previous findings that accounting ratios and macroeconomic variables play distinct roles in default prediction for private firms: (i) Accounting ratios are necessary for accurately ranking private firms according to default likelihood but cannot by themselves capture the cyclicality of aggregate default rates, while (ii) macroeconomic variables are indispensable for capturing the cyclicality of aggregate default rates but do not aid in the ranking of firms with respect to default likelihood. Using our fitted Cox model, we find that when we split our sample with respect to firm size and keep all other firm characteristics fixed, the default probabilities of smaller firms do in fact exhibit less sensitivity to macroeconomic cycles. This is in the sense that the effects of macroeconomic variables are generally of smaller magnitude for smaller firms. However, when we account for the Cox model's non-linear form and use averaging techniques adapted from other non-linear regression models, our results indicate that the default probability of the average small firm may be as cyclical as or even more cyclical than the default probability of the average large firm.

Second, we investigate our data using the more flexible additive Aalen regression model. The additive model allows us to estimate time-varying effects of firm-specific variables, which is a potentially important component of the cyclicality of default probabilities that is missing from our Cox model. Furthermore, due to the linearity of its effects, the additive model allows us to directly compare the effects of macroeconomic variables for small and large firms without the need for averaging techniques. Our analysis based on the additive model reveals, in particular, that firm size has a significantly time-varying and mostly negative effect. Hence, the effect of firm size varies significantly with the business cycle but is generally such that larger firms have lower default probabilities. Moreover, the effect of firm size reaches its largest (negative) magnitude during the financial crisis of 2008–10, which indicates that larger firms were safer, not riskier, during the most recent recession. Lastly, using our fitted additive model and splitting the sample with respect to firm size, we again find no evidence that there is a difference in the signs, magnitudes, or significance of the effects of our macroeconomic variables for small firms compared with large firms. These results indicate that our main findings are robust to different regression models that capture cyclicality in different ways.

Our results indicate that a different treatment of capital charges solely based on firm size is too simplistic, as it ignores other important characteristics that differ between small and large firms. This suggests that imposing solely size-based preferential treatment of capital charges may, in fact, increase the risks of banks with a high exposure to the SME segment.

The outline of the paper is as follows. Section 2 explains the piece of the Basel regulation that motivates our study. Section 3 reviews the literature. Section 4 details our data and variable selection. Section 5 provides our estimation methodology. Section 6 presents our regression results for identifying the firm-specific and macroeconomic variables that significantly predict defaults. Section 6 presents our results regarding the sensitivity of small and large firms' default probabilities to macroeconomic variables. Section 7 shows results related to robustness and model check. Section 8 concludes.

2. The Basel Capital Charge for Loans

Basel II and III allow banks to estimate capital requirements for small and medium-sized corporations (SMEs) using a risk-weight formula that includes a lower asset correlation with macroeconomic risk drivers compared with that of larger corporations. The exact formulation in Basel II can be found in articles 273 and 274 of the Basel Committee on Banking Supervision's (2006) report. The same provisions are carried forward and extended in the Basel III Accord and the recently adopted CRD IV as articles 153.4 and 501.1 of the European Parliament and the Council of the European Union's (2013) report.

The amount of equity capital required for funding a loan essentially depends on the unexpected loss (UL) per unit principal,

$$\mathrm{UL} = \mathrm{LGD}(\theta^* - \mathrm{PD}),$$

where PD is the loan's one-year probability of default, LGD is the loss rate given default, and

$$\theta^* = \Phi\left(\frac{1}{\sqrt{1-\rho}}\left(\Phi^{-1}(\mathrm{PD}) + \sqrt{\rho}\,\Phi^{-1}(0.999)\right)\right). \tag{1}$$

Here, Φ is the standard normal distribution function and θ^* is the 99.9 percent worst-case default rate. The correlation parameter, ρ, that goes into this calculation is PD dependent and prescribed by the Basel documentation (after removing the negligible terms $\exp(-50)$) as

$$\rho(\mathrm{PD}) = 0.12\,(1 - \exp(-50\,\mathrm{PD})) + 0.24\exp(-50\,\mathrm{PD}).$$

For SMEs with annual sales, S, up to €50 million, this correlation is modified to

$$\begin{aligned}\rho(\mathrm{PD}) &= 0.12\,(1 - \exp(-50\,\mathrm{PD})) + 0.24\exp(-50\,\mathrm{PD}) \\ &\quad - 0.04\left(1 - \frac{\max(S-5;0)}{45}\right),\end{aligned}$$

where the last term has the effect of lowering the capital requirement. The exact capital requirement depends on the contribution to risk-weighted assets (RWA), which depends on UL, the exposure at default (EAD), and a maturity adjustment whose exact form need not concern us here.

As an illustration of the economic magnitude of the reduction in capital requirement for SMEs, note that with an annual default probability of 1 percent, a loan to an SME with S = €25 million achieves a deduction in asset correlation of 11.52 percent relative to an equally risky non-SME. Assuming an LGD of 15 percent and an effective maturity of 2.5 years (both provided as "representative averages" by the financial institution that provided us with the data), we calculate that this corresponds to a deduction of 12.14 percent in capital requirement.

In the derivation of the capital charge formula, ρ represents the correlation between asset values of different borrowers that arises because the asset values depend on a common economy-wide factor. Therefore, the reduction in correlation for SMEs corresponds to an assumption that these firms have default probabilities that are less sensitive to the economy-wide factor. It is this assumption that we will devise two ways of testing on a large Danish data set.

3. Related Literature

There is conflicting evidence in the literature regarding the validity of the assumption in the Basel Accords that the default probabilities of smaller firms are less cyclical. Lopez (2004) employs the KMV approach on a sample of U.S., Japanese, and European firms and finds that average asset correlation is a decreasing function of probability of default and an increasing function of firm size, which is line with the assumption in the Basel Accords. Dietsch and Petey (2004), however, use a one-factor credit risk model on a sample of French and German firms and find that SMEs have higher default risk than larger firms and that the asset correlations for SMEs are weak and decrease with firm size. Chionsini, Marcucci, and Quagliariello (2010) use firm-size dependent linear regression of default probabilities on macroeconomic time series for a sample of Italian firms, and they find support for the size-dependent treatment in the Basel Accord, though not during severe financial crises like that of 2008–09. Jacobson, Lindé, and Roszbach (2005) use a Monte Carlo resampling method for a sample of Swedish firms and find little support for the hypothesis that SME loan portfolios are less risky, or require less economic capital, than corporate loans.

Our approach differs from these studies in that we use an intensity regression framework to directly estimate and study how default probabilities simultaneously depend on a large set of firm characteristics, including size, as well as macroeconomic variables. This allows us to determine the firm characteristics that are important for accurately ranking firms with respect to default likelihood as well as the macroeconomic variables important for capturing the cyclicality of default rates over time. Importantly, our approach has the advantage of allowing us to directly determine how the effects of macroeconomic variables on default probabilities differ for small and large firms while simultaneously controlling for firm characteristics other than size. We apply averaging techniques to the Cox regression model resembling those commonly applied to generalized linear models—see, for instance, Wooldridge (2009, 582–83) for an overview. The additive regression model due to Aalen (1980, 1989) was applied in a default prediction setting by Lando et al. (2013). They show that there is significant time variation in how certain firm-specific variables influence the default probabilities of public U.S. firms.

Statistical models using accounting ratios to estimate default probabilities date back to at least Beaver (1966) and Altman (1968), followed by Ohlsen (1980) and Zmijewski (1984)—we use many of the same accounting ratios as in these studies. Shumway (2001) was among the first to demonstrate the advantages of intensity models with time-varying covariates compared with traditional discriminant analysis, and was also among the first to include equity return as a market-based predictor of default probabilities—we use a similar estimation setup, although we do not have market-based variables for our private firms. Chava and Jarrow (2004) improved the setup of Shumway (2001) using covariates measured at the monthly level and showed the importance of industry effects—our data frequency is also at the monthly level and we correct for industry effects in all our regressions.

Structural models of credit risk, such as the models of Black and Scholes (1973), Merton (1974), and Leland (1994), usually assume that a firm defaults when its assets drop to a sufficiently low level relative to its liabilities. The connection between structural models and intensity models was formally established by Duffie and Lando (2001), who showed that when the firm's asset value process is not perfectly observable, a firm's default time has a default intensity that depends on the firm's observable characteristics as well as other covariates. Studies demonstrating the importance of covariates implied from structural models, such as distance to default or asset volatility, include Duffie, Saita, and Wang (2007), Bharath and Shumway (2008), Lando and Nielsen (2010), and Chava, Stefanescu, and Turnbull (2011), among many others.

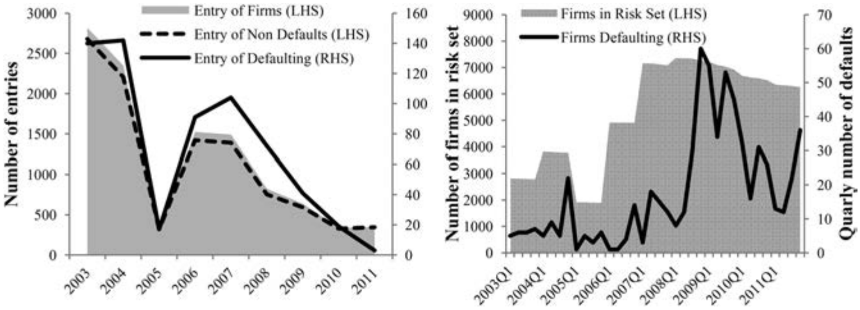
Default studies using data on public firms and demonstrating the importance of employing macroeconomic variables include McDonald and de Gucht (1999), Pesaran et al. (2006), Duffie, Saita, and Wang (2007), Lando and Nielsen (2010), and Figlewski, Frydman, and Liang (2012), among many others. Recent default studies of private firms that also employ macroeconomic variables include Carling et al. (2007), who use Swedish data, and Bonfim (2009), who uses Portuguese data. We employ many of the same macroeconomic variables as in these studies.

4. Data and Variables

This section presents our data and the variables that we employ as firm-specific and macroeconomic drivers of default probabilities.

Our raw data comprise 28,395 firms and 114,409 firm-year observations of obligor and loan histories, accounting statements, and default indicators over the period 2003 to 2012. The data are obtained from a large Danish A-IRB (advanced internal ratings-based approach) financial institution. A firm is included in this data set if it has an engagement over DKK 2 million in at least one of the years underlying the period of analysis. An engagement is defined in terms of loans or granted credit lines. After removing sole proprietorships, government institutions, holding companies without consolidated financial statements, firms that do not have Denmark as their residency, and firms with insufficient balance sheet information, we are left with 10,671 firms and 48,703 firm-year observations. In the cleaned data set, a total of 633 firms experienced a default event, defined by the Basel II Accord as more than ninety days delinquency. Moreover, 54 of the 633 defaulting firms experience a second default, in the sense that they became delinquent a second time during the sample period. Other default studies have treated a firm that reemerges from default as a new firm. In accordance with the Basel II Accord's definition of a default event as a period of delinquency, we choose to disregard multiple default events, so that only the initial default counts.

Figure 1. Entry and At-Risk Pattern in the Sample

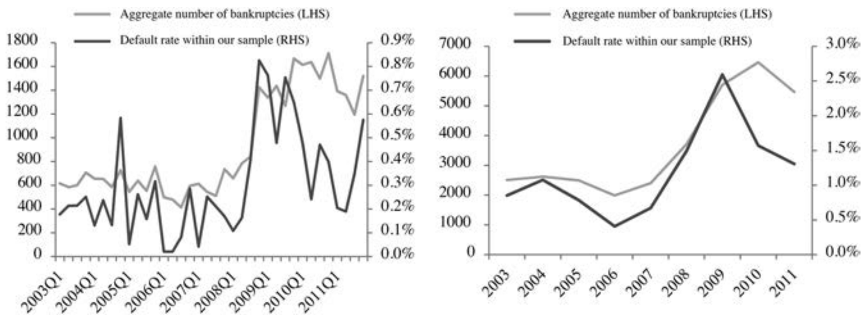


Notes: The left panel shows the yearly number of firms entering the sample (grey shaded area) along with the yearly number of entries that do not default (solid black line) or eventually do default (dashed black line). The right panel shows the quarterly number of firms at risk of defaulting (i.e., in the "risk set"; grey shaded area) along with the actual number of defaulting firms in each quarter (solid black line).

Figure 1 shows the patterns by which firms enter and potentially leave our final sample. The left panel shows the number of firms that enter the sample at each year along with an indication of the number of entries eventually corresponding to defaults and non-defaults. Despite discussions with the financial institution providing the data, the low number of firms entering the sample in 2005 remains a conundrum. It appears, however, that the firms that eventually default do not seem to differ systematically from the non-defaulting firms based on when they enter the sample. The right panel shows the number of firms at risk of defaulting, i.e., firms in the "risk set," at each quarter, along with the quarterly number of defaults. The risk set is seen to contain at least 2,000 firms at each quarter, and the 2008–09 financial crisis is readily visible from the sharp rise in the number of defaults.

In order to incorporate quarterly macroeconomic variables, we recode the accounting variables for each firm from annual to quarterly observations. This will naturally induce persistence in the accounting variables from quarter to quarter, which we correct for by basing all inference on standard errors clustered at the firm level. The final data set thus consists of a total of 192,196 firm-quarter observations.

Figure 2. Default Rates in the Sample and the General Danish Economy



Notes: The left panel shows the quarterly default rate from the sample along with the corresponding aggregate number of quarterly bankruptcies in the general Danish economy. The right panel shows the yearly default rate from the sample along with the aggregate quarterly number of defaults in the general Danish economy.

Figure 2 compares the observed default rate in our sample with the number of registered bankruptcies in Denmark. The comparison is feasible because the total number of firms at risk of default in Denmark is relatively stable over time. We see that, due to the relatively few incidences, the default rate in the sample co-moves nicely with the aggregate level in Denmark. This indicates that our results are likely to be representative for all Danish SMEs.

4.1 Firm-Specific Explanatory Variables

Table 1 provides an overview of the firm-specific explanatory variables which we employ in our regression analysis. Our firm-specific explanatory variables measure size, age, leverage, profitability, asset liquidity, collateralization, and (book) equity. The table also gives the accounting ratios that we use to proxy for the firm-specific explanatory variables, along with the expected sign of each accounting-based variable's effect on default probabilities. All our accounting-based variables have been applied in previous default studies, and the expected signs of their effects are both intuitive and well discussed in the literature—see, for instance, Ohlsen (1980),

Table 1. Firm-Specific Explanatory Variables and Corresponding, Observable Accounting-Based Variables

Explanatory Variable	Proxy	Expected Effect on Probability of Default
Size	Log of Book Value of Assets	Negative
Age	Years Active in the Bank	Negative
Leverage	Short-Term Debt to Total Assets	Positive
	Total Debt to Total Assets	Positive
	Interest-Bearing Debt to Total Assets	Positive
	Interest Payments to Total Assets	Positive
Profitability	Net Income to Total Assets	Negative
	EBIT to Total Assets	Negative
	EITDA to Total Assets	Negative
Liquidity	Current Ratio	Negative
	Quick Ratio	Negative
Collateralization	Fixed Assets to Total Assets	Negative
	PPE to Total Assets	Negative
Negative Equity	Dummy for Negative Equity	Positive
Industry Effects	Sector Affiliation	Control Variable

Notes: The left column shows our list of firm-specific explanatory variables, the center column shows the observable accounting-based variables that we use as proxies, and the right column showa the expected effect of each acounting-based variable on default probabilities. Industry effects are included in the list for completeness, although we only use this variable as a control (see details in the text).

Shumway (2001), Duffie, Saita, and Wang (2007), and Lando and Nielsen (2010). We also correct for industry effects as in Chava and Jarrow (2004). The main difference between our list of firm-specific variables and the ones used in default studies of public firms is the lack of market-based measures such as stock return and distance to default.

We control for industry effects since certain industry characteristics may prescribe a certain leverage structure—for example, because of differences in the volatility of cash flows. We use the sector affiliation by Statistics Denmark to identify a firm's primary industry as either "construction," "manufacturing," or "wholesale and retail," as these have above-average default rates but are at the same time coarse enough to ensure a sufficient number of firms in each sector.

Table 2 presents summary statistics for our firm-specific explanatory variables. An analysis of the variables revealed a few miscodings and extreme values. Due to the anonymized nature of the data, we were not able to check the validity of these data points manually, and we therefore choose to winsorize all the firm-specific variables at the 1st and 99th percentile—a practice also used by Shumway (2001), Chava and Jarrow (2004), and Bonfim (2009), among others. The average firm has DKK 275 million in assets, a total-debt-to-total-assets ratio of 68 percent, and interest payments corresponding to 3 percent of total assets. Further, the average firm had a relationship with the bank for twenty-three years and remains in the sample for seven out of the nine years.

Due to Danish reporting standards, firms below a certain size may refrain from reporting revenue and employee count, and hence these variables are zero (or missing) for a large proportion of firms in the sample. We therefore choose not to use these two variables in our further analysis in order to retain a large sample of smaller firms. In table 2, firm age is taken to be the time since the bank recorded the first interaction with the client, entry year specifies the year at which the firm enters the sample, and duration is the number of years a firm is observed in the sample since its entry year. The negative (book) equity dummy has an unconditional mean of 0.063, meaning that just over 6 percent of our firm-quarters show a negative value of (book) equity.

4.2 Macroeconomic Explanatory Variables

Table 3 provides an overview of the macroeconomic explanatory variables that we employ in our regression analysis. Our macroeconomic explanatory variables cover the stock market, interest rates, GDP, credit supply, inflation, industrial production, and demand for consumer goods. The table also gives the observable time series that we use to proxy for the macroeconomic explanatory variables, along with the expected sign of each time series' effect on default probabilities. The macroeconomic time series are primarily obtained from

Table 2. Descriptive Statistics for the Firm-Specific Variables

Variable	Mean	Std.	1%	5%	25%	50%	75%	95%	99%
Total Assets (tDKK)	275,074	1,060,373	811	2,784	9,553	28,222	100,723	1,025,875	8,656,000
Revenue (tDKK)	242,031	885,431	0	0	0	0	70,925	1,097,486	6,733,409
Employees	113	348	0	0	1	18	64	484	2,666
Age (Years)	23	20	1	3	9	18	30	71	97.25
Log(Total Assets) (tDKK)	10.46	1.81	6.70	7.93	9.16	10.25	11.52	13.84	15.97
Short-Term Debt to Total Assets	0.51	0.28	0.01	0.09	0.30	0.49	0.69	0.96	1.58
Total Debt to Total Assets	0.68	0.28	0.02	0.18	0.53	0.70	0.84	1.04	1.80
Interest-Bearing Debt to Total Assets	0.39	0.28	0.00	0.00	0.17	0.37	0.56	0.87	1.38
Interest Payments to Total Assets	0.03	0.03	0.00	0.00	0.01	0.02	0.03	0.07	0.17
Current Ratio	1.66	2.42	0.03	0.28	0.88	1.17	1.59	3.80	20.21
Quick Ratio	1.28	2.37	0.02	0.14	0.49	0.81	1.19	3.15	19.67
Fixed Assets to Total Assets	0.40	0.29	0.00	0.01	0.14	0.36	0.63	0.93	0.99
Tangible Assets to Total Assets	0.30	0.28	0.00	0.00	0.06	0.23	0.49	0.87	0.97
Net Income to Total Assets	0.03	0.14	-0.68	-0.18	0.00	0.03	0.09	0.24	0.43
EBIT to Total Assets	0.06	0.15	-0.59	-0.16	0.00	0.05	0.12	0.29	0.51
EBITDA to Total Assets	0.10	0.15	-0.51	-0.12	0.02	0.09	0.17	0.34	0.54
Entry Year	2005	1.94	2003	2003	2003	2004	2006	2009	2010
Duration (Years)	7	2	1	2	5	7	9	9	9

Notes: The table shows descriptive statistics for the firm-specific variables of the cleaned sample, winsorized at the 1st and 99th percentile. The total number of observations is 192,196 firm-quarters. Age is time since the bank recorded the first interaction with the client. Entry is the year in which the firm entered the sample. Duration is the number of years the firm remains in the sample. All other variables have standard interpretations.

Table 3. Macroeconomic Explanatory Variables and Corresponding Observable Time Series

Explanatory Variable	Proxy	Expected Effect on Probability of Default
Stock Return	Return of OMX Index	Negative
Stock Volatility	Volatility of OMX Index	Unknown
Interest Rates	Slope of Yield Curve	Negative
GDP	Real Growth in Danish GDP	Negative
Loan Growth	Loan Growth to Non-financial Firms	Positive
Credit Availability	Funding Costs	Positive
Aggregate Defaults	Danish Bankruptcies	Positive
Inflation	Headline CPI	Unknown
Demand-Side Effects	Consumer Confidence	Negative
	House Prices	Negative
Supply-Side Effects	Business Indicator, Manufacturing	Negative
	Capacity Utilization	Negative
International Exposure	Exports to Danish GDP	Unknown
	Return of Stoxx50 Index	Negative
	EU-27 GDP Growth	Negative

Notes: The left column shows our list of macroeconomic explanatory variables, the center column shows the macroeconomic time series that we use to measure each macroeconomic explanatory variable, and the right column shows the expected effect of each macroeconomic time series on default probabilities.

Ecowin, with additional data from Statistics Denmark, the OECD, and Stoxx.

The inclusion of lagged macroeconomic variables allows us to use these to compute growth rates, differences, or levels. We select the appropriate form by (i) computing the correlation between each form of the macroeconomic variable and the observed default rate, and (ii) visually inspecting the relationship of each form with the observed default rate. Note, however, that some pairs of the macroeconomic variables exhibit collinearity—for example, the Danish GDP growth and the European GDP growth rate, as well as the return on the

OMX index and the Stoxx index, have pairwise correlations of 0,92 and 0,77, respectively. The high degree of collinearity should be kept in mind when interpreting the estimated regression coefficients in the following sections.

5. Estimation Methodology

This section presents the methodology that we apply to estimate the effects of our explanatory variables on firm-specific default probabilities.

Suppose we have a sample of n levered firms observed over a time horizon $[0, T]$, where firm i may default at a stochastic time τ_i. At each time t, the firm's financial state is determined by a vector $\mathbf{X}_{it}$ of firm-specific variables and by a vector $\mathbf{Z}_t$ of macroeconomic variables, with values common to all firms in the sample. Default at time t occurs with intensity $\lambda_{it} = \lambda(\mathbf{X}_{it}, \mathbf{Z}_t)$, meaning that λ_{it} is the conditional mean arrival rate of default for firm i, measured in events per time unit. Intuitively, this means that, given survival and the observed covariate histories up to time t, firm i defaults in the short time interval $[t, t + dt)$ with probability $\lambda_{it}\, dt$.[2] We assume τ_i is doubly stochastic driven by the combined history of the internal and external covariates (see, for instance, Duffie, Saita, and Wang 2007).

We devise two different ways of testing whether defaults of smaller firms are more or less cyclical. First, we use the standard multiplicative Cox regression model to estimate the effects of firm-specific and macroeconomic variables on default intensities. Second, we use the more flexible additive Aalen regression model, which allows for time-varying effects of firm-specific variables.

5.1 The Cox Regression Model

In our initial analysis of which accounting ratios and macroeconomic variables significantly predict defaults, we specify the firm-specific

[2]Precisely, a martingale is defined by $1_{(\tau_i \leq t)} - \int_0^t 1_{(\tau_i > s)} \lambda_{is}\, ds$ with respect to the filtration generated by the event $(\tau_i > t)$ and the combined history of the firm-specific and macroeconomic variables up to time t.

default intensities using the popular "proportional hazards" regression model of Cox (1972). The intensity of firm i at time t is thus modeled as

$$\lambda(\mathbf{X}_{it}, \mathbf{Z}_t) = Y_{it} \exp\left(\boldsymbol{\beta}^\top \mathbf{X}_{it} + \boldsymbol{\gamma}^\top \mathbf{Z}_t\right),$$

where Y_{it} is an at-risk indicator for firm i, taking the value of 1 if firm i has not defaulted "just before" time t and 0 otherwise, while $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are vectors of regression coefficients. The effect of a one-unit increase in the jth internal covariate at time t is to multiply the intensity by the "relative risk" e^{β_j}. The same interpretation applies to the external covariates. We let the first component of the vector $\mathbf{Z}_t$ be a constant 1. This means that the first component of $\boldsymbol{\gamma}$ is a baseline intensity, corresponding to the (artificial) default intensity of firm i when all observable covariates are identically zero.[3]

Following, for instance, Andersen et al. (1992), and under the standard assumptions that late-entry, temporal withdrawal, right-censoring, and covariate distributions are uninformative on regression coefficients, the (partial) log-likehood for estimation of the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ based on a sample of n firms becomes

$$\begin{aligned} l(\boldsymbol{\beta}, \boldsymbol{\gamma}) = & \sum_{i=1}^{n} \int_0^T \left(\boldsymbol{\beta}^\top \mathbf{X}_{it} + \boldsymbol{\gamma}^\top \mathbf{Z}_t\right) dN_{it} \\ & - \int_0^T \sum_{i=1}^{n} Y_{it} \exp\left(\boldsymbol{\beta}^\top \mathbf{X}_{it} + \boldsymbol{\gamma}^\top \mathbf{Z}_t\right) dt, \end{aligned}$$

where $N_{it} = 1_{(\tau_i \leq t)}$ is the one-jump default counting process for firm i. We investigate the assumption of independent censoring and entry pattern in section 8, and find that our parameter estimates are robust to the exclusion of firm-years that could potentially induce bias.

Estimation, inference, and model selection for the Cox model may then be based on maximum-likelihood techniques. Given

[3]Note that while the usual Cox model includes an (unspecified) time-varying baseline intensity, thereby making it a semi-parametric survival regression model, we cannot simultaneously identify the vector $\boldsymbol{\gamma}$ of macroeconomic regression coefficients as well as a time-varying baseline intensity—we therefore restrict to a fully parametric model with a constant baseline intensity.

maximum-likelihood estimators (MLEs) of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, we can judge the influence of covariates on default intensities by judging the significance of the corresponding regression coefficients, and we can predict firm-specific and aggregate default intensities by plugging the MLEs back into intensity specification of the Cox model. A model check may be based on the so-called martingale residual processes,

$$N_{it} - \int_0^t Y_{is} \exp\left(\widehat{\boldsymbol{\beta}}^\top \mathbf{X}_{is} + \widehat{\boldsymbol{\gamma}}^\top \mathbf{Z}_s\right) ds, \quad i = 1, \ldots, n, \quad t \in [0, T], \tag{2}$$

which, when the model is correct, converge to mean-zero martingales as the sample size increases. Hence, when aggregated over covariate-quantiles or sectors, the grouped residuals processes should not exhibit any systematic trends when plotted as functions of time.

5.2 The Aalen Regression Model

In addition to the Cox model, we will also employ the additive regression model of Aalen (1980, 1989), which specifies the default intensity of firm i as a linear function of the covariates. This allows us to estimate time-varying effects of firm-specific variables, which is a potentially important component of the cyclicality of default probabilities that is missing from our Cox models. Furthermore, due to the linearity of its effects, the additive model allows us to directly compare the effects of macroeconomic variables for small and large firms without the need for averaging techniques, in contrast to the Cox model.

Our specification of the additive model for the default intensity of firm i is given by

$$\lambda(\mathbf{X}_{it}, \mathbf{Z}_t) = \boldsymbol{\beta}(t)^\top \mathbf{X}_{it} + \boldsymbol{\gamma}^\top \mathbf{Z}_t,$$

where $\boldsymbol{\beta}(t)$ is a vector of unspecified regression *functions of time*, while $\boldsymbol{\gamma}$ is a vector of (time-constant) regression coefficients.[4]

[4]Note that we cannot identify time-varying effects for the macroeconomic variables, as the macroeconomic variables do not vary across firms at a given point in time.

The linearity of the additive model allows for estimation of both time-varying and constant parameters using ordinary least squares methods. For the time-varying coefficients, the focus is on the *cumulative regression coefficients*, $B_j(t) = \int_0^t \beta_j(s)\, ds$, which are easy to estimate non-parametrically. Further, formal tests of the significance and time variation of regression functions are possible through resampling schemes. We refer to Martinussen and Scheike (2006), Aalen, Borgan, and Gjessing (2008), and Lando et al. (2013) for a detailed presentation of estimation and inference procedures.

6. Default Prediction for Private Firms

In this section, we investigate which accounting ratios and macroeconomic variables significantly predict defaults in our sample. We initially focus on the multiplicative Cox model, before using the additive Aalen model to judge whether allowing for time-varying effects of firm-specific variables changes our conclusions.[5]

First, we show the result from a Cox model using only firm-specific variables. We will see that this model cannot adequately predict the cyclical variation in the aggregate default rate. Second, we add macroeconomic variables to the Cox model and show that this allows the model to much more accurately predict the aggregate default rate over time. However, when judging the different Cox models' ability to correctly rank firms with respect to default likelihood, we will see that macroeconomic variables only marginally improve the ranking based on accounting ratios alone. Hence, to capture the cyclicality of default rates, it is sufficient to focus on macroeconomic variables—however, accounting variables

[5]We focus on both the Cox and the Aalen models because they each have their advantages in the tests we consider. The Cox model has the advantages that it automatically produces non-negative intensities and that its constant regression coefficients allow out-of-sample prediction. The Aalen model has the advantages of time-varying effects of firm-specific variables and easy comparison of effects across subsamples. Note that while there exist variants of the Cox model that in principle allow for time-varying effects of firm-specific variables, such variants have to apply some degree of smoothing in each estimation iteration, which might blur or distort the effects. Such smoothing is not necessary for estimating time-varying effects in the additive Aalen model. See Martinussen and Scheike (2006), Aalen, Borgan, and Gjessing (2008), and Lando et al. (2013) for more on the differences between (variants of) the Cox model and the Aalen model.

are necessary controls for variations in firm-specific default risk not related to size. Finally, we estimate an additive Aalen using the same variables as in our preferred Cox model, but allowing for time-varying effects of the firm-specific variables. We find significant time variation in the effects of several firm-specific variables. Furthermore, allowing for this time variation turns the effects of the macroeconomic variables insignificant. Hence, time variation in the effects of firm-specific variables is an important alternative way of quantifying the cyclicality of default probabilities.

6.1 Using Accounting Ratios Alone

Initially, we fit a Cox model of firm-by-firm default intensities using only firm-specific variables. We will use this fitted model to examine to what extent macroeconomic variables add additional explanatory power to default prediction.

Table 4 presents estimation results for Cox models using only firm-specific variables. Due to the high degree of correlation among the measurements within the same categories, we perform a stepwise elimination of variables in a given category, removing the least significant variables in each step. The outcome is that interest-bearing debt to total assets, net income to total assets, the quick ratio, and tangible assets (PPE) to total assets remain in the model, along with age of banking relationship, log of total assets, and a negative equity dummy.

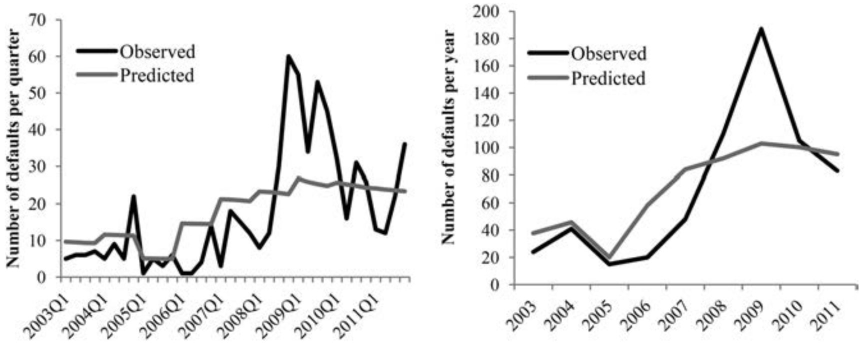
Interpreting the preferred model (model 5 in table 4), the effect of age is negative, implying that the longer a firm has had a relationship with the bank, the less likely it is that the firm will default. The effect of size, as measured by book assets, appears insignificant in the specification. This might potentially be explained by the sample pertaining to only the largest corporate clients, where size is less relevant as an explanation of default. The leverage ratio of interest-bearing debt to total assets is, as expected, positively related to default probability. Likewise, past profitability is negatively related to default probability. The quick ratio enters with a significant negative sign, confirming the hypothesis that the more liquidity a firm has, the higher its ability to service unexpected cash shortfalls that would otherwise have resulted in a default. Tangible assets, measured as plant property and equipment (PPE) to total assets, does

Table 4. Estimation Results for Cox Models including Only Accounting Ratios

Dependent Variable: Default (0/1)	Model 1	Model 2	Model 3	Model 4	Model 5
Variables (All Lagged One Year)	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**
Intercept	-6.788^{***}	-6.494^{***}	-6.668^{***}	-6.684^{***}	-6.729^{***}
Years Active in the Bank	-0.011^{***}	-0.011^{***}	-0.011^{***}	-0.011^{***}	-0.011^{***}
Log of Total Assets	0.043^{*}	0.014	0.020	0.019	0.016
(i) Short-Term Debt to Total Assets	-0.572^{**}				
Total Debt to Total Assets	0.644^{**}				
Interest-Bearing Debt to Total Assets	0.661^{***}	1.226^{***}	1.262^{***}	1.261^{***}	1.255^{***}
Interest Payments to Total Assets	7.843^{***}				
(ii) Net Income to Total Assets	-0.961^{**}	-1.790^{***}	-2.012^{***}	-2.025^{***}	-2.015^{***}
EBIT to Total Assets	1.403^{*}	1.963^{**}			
EBITDA to Total Assets	-2.615^{***}	-2.514^{***}			
(iii) Quick Ratio	−0.078	−0.045	−0.057	-0.212^{**}	-0.202^{**}
Current Ratio	−0.192	−0.179	−0.161		
(iv) Fixed Assets to Total Assets	−0.455	−0.288	−0.308	−0.280	
PPE to Total Assets	0.529^{**}	0.571^{**}	0.473^{**}	0.466^{**}	0.255
Negative Equity, Dummy	0.467^{**}	0.547^{***}	0.555^{***}	0.563^{***}	0.570^{***}
Construction, Dummy	0.926^{***}	0.885^{***}	0.928^{***}	0.921^{***}	0.951^{***}
Wholesale and Retail Trade, Dummy	0.214	0.216^{*}	0.267^{**}	0.250^{*}	0.275^{**}
Manufacturing, Dummy	0.404^{***}	0.399^{***}	0.420^{***}	0.406^{***}	0.417^{***}
Number of Observations	192,196	192,196	192,196	192,196	192,196
Number of Firms	10,671	10,671	10,671	10,671	10,671
Number of Events	633	633	633	633	633
Sector Effects	Yes	Yes	Yes	Yes	Yes
QIC	7,677.2	7,716.4	7,724.8	7,722.8	7,721.9
QICu	7,669.2	7,710.5	7,719.2	7,717.8	7,717.5

Notes: The table shows parameter estimates, standard errors, and model summary statistics for Cox models of the quarterly default intensity of firms in the sample. All variables are lagged one year to allow for one-year prediction. The full list of firm-specific variables is included in model 1. Models 2–4 show the stepwise elimination, keeping only the most significant measure within the groups of (i) leverage, (ii) profitability, (iii) liquidity, and (iv) collateralization. Model 5 (shaded grey) is the preferred specification when only firm-specific variables are used as covariates. *, **, and *** denote parameter significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Parameter significance is based on robust standard errors clustered at the firm level.

Figure 3. Default Prediction Based on the Preferred Cox Model only Including Accounting Ratios



Notes: The left panel shows the observed number of quarterly defaults in the sample along with the predicted number of defaults based on the preferred Cox model only including accounting ratios (model 5 in table 4). The right panel is similar, except that the aggregation is done on a yearly basis.

not appear to have a significant effect, confirming the findings of Bonfim (2009) that tangible assets remain insignificant in explaining corporate defaults. The negative (book) equity dummy enters with a positive sign in all specifications, confirming that negative (book) equity is in fact a sign of a firm in trouble and at increased risk of default. The sign of the sectoral dummies are all positive and significant, confirming that these sectors have above-average default rates.

Using the results of table 4, we calculate a predicted quarterly default intensity for each firm in the sample, and then aggregate these to get a predicted aggregate intensity for each quarter. Figure 3 shows the observed number of quarterly defaults in the sample, along with the predicted number of defaults based on the preferred Cox model only including accounting ratios. As is evident in the left panel, the model based on accounting ratios alone is unable to explain the cyclical nature of the observed defaults. However, acknowledging that the firm-specific data can only change yearly through annual financial statements, it may be more appropriate to aggregate the predicted and observed number of defaults on a yearly basis. This is shown in the right panel, and the conclusion is the same: The model based solely on firm-specific variables is not capable of capturing the cyclical variation in defaults.

6.2 Including Macroeconomic Variables

Given that firm-specific variables are unable to explain the cyclical nature of defaults in our sample, this section incorporates macroeconomic effects in our Cox model. In order to assess if macroeconomic variables add explanatory power beyond what is implied by the firm-specific variables, the preferred model of the firm-specific variables is used as the basis of the covariate specification.

Table 5 presents estimation results for Cox models incorporating macroeconomic variables. The selection procedure has been to perform a stepwise elimination of insignificant variables until only significant macroeconomic variables remain in the model. Model 7 is the preferred model including both firm-specific and microeconomic variables, while model 8 is this preferred model excluding the firm-specific variables.

The effects of the firm-specific variables remain robust to the inclusion of the macroeconomic variables. In the preferred model (model 7 of table 5), the significant macroeconomic variables are as follows: the return of the OMX stock market index, the volatility of the OMX index, the difference between CIBOR and the policy rate, the slope of the yield curve, the change in consumer confidence, the change in the capacity utilization, the return of the Stoxx 50 index, and, finally, the European GDP growth rate. On the other hand, the aggregate number of defaults, the Danish real GDP growth, exports as a fraction of GDP, inflation, changes in the cyclical indicator for construction, and the loan growth to non-financials are all insignificant.

When interpreting the coefficients of the macroeconomic variables in multivariate intensity regression models, one should bear in mind that it would be unrealistic to obtain a complete ceteris paribus effect of one macroeconomic variable, as this variable cannot be viewed in isolation from other macroeconomic variables. While not done here, an appropriate interpretation would involve testing the model from the perspective of internally consistent scenarios of macroeconomic variables. For instance, a further analysis shows that the volatility of the stock market, the slope of the yield curve, the

Table 5. Estimation Results for Cox Models with both Accounting and Macroeconomic Variables

Dependent Variable: (0/1)	**Model 6**	**Model 7**	**Model 8**
Variables (All Lagged One Year)	**Coef.**	**Coef.**	**Coef.**
Intercept	−13.341***	−8.051***	−7.206***
Years Active in the Bank	−0.011***	−0.011***	
Log of Total Assets	0.007	0.007	
Interest-Bearing Debt to Total Assets	1.232***	1.231***	
Net Income to Total Assets	−1.877***	−1.877***	
Quick Ratio	−0.200**	−0.200**	
PPE to Total Assets	0.282*	0.281*	
Dummy for Negative Equity	0.592***	0.591***	
Aggregate Quarterly Number of Danish Bankruptcies	0.005		
Danish Real GDP Growth	−0.027		
Export/GDP	9.633*		
Inflation, Pct. Point	−0.124		
OMX Stock Market Return	−0.052***	−0.045***	−0.038***
OMX Stock Market Volatility	−0.119***	−0.101***	−0.099***
Difference CIBOR – Policy Rate, Pct. Point	1.837***	2.006***	2.117***
Yield-Curve Slope 10y – 3m, Pct. Point	0.986***	0.588***	0.634***
Growth in House Prices	−0.114***	−0.104***	−0.115***
Change in Consumer Confidence Indicator	−0.110***	−0.099***	−0.110***
Change in Cyclical Indicator, Construction	0.002		
Change in Capacity Utilization in the Industrial Sector	−0.338***	−0.300***	−0.292***
Loan Growth to Non-financial Institutions	0.023**		
Stoxx50 Stock Market Return	0.031***	0.034***	0.029***
EU-27 Real GDP Growth	0.250***	0.227***	0.213***
Number of Observations	192,196	192,196	192,196
Number of Firms	10,671	10,671	10,671
Number of Events	633	633	633
Sector Effects	Yes	Yes	Yes
QIC	7,513	7,513.6	8,265.0
QICu	7,518	7,508.6	8,264.7

Notes: The table shows parameter estimates, standard errors, and model summary statistics for Cox models of the quarterly default intensity of firms in the sample. All variables are lagged one year to allow for one-year prediction. The full list of firm-specific and macroeconomic variables is included in model 6, and model 7 (shaded grey) is the preferred specification after stepwise elimination of variables. Model 8 is the preferred specification in model 7 excluding the firm-specific variables. *, **, and *** denote parameter significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Parameter significance is based on standard errors clustered at the firm level.

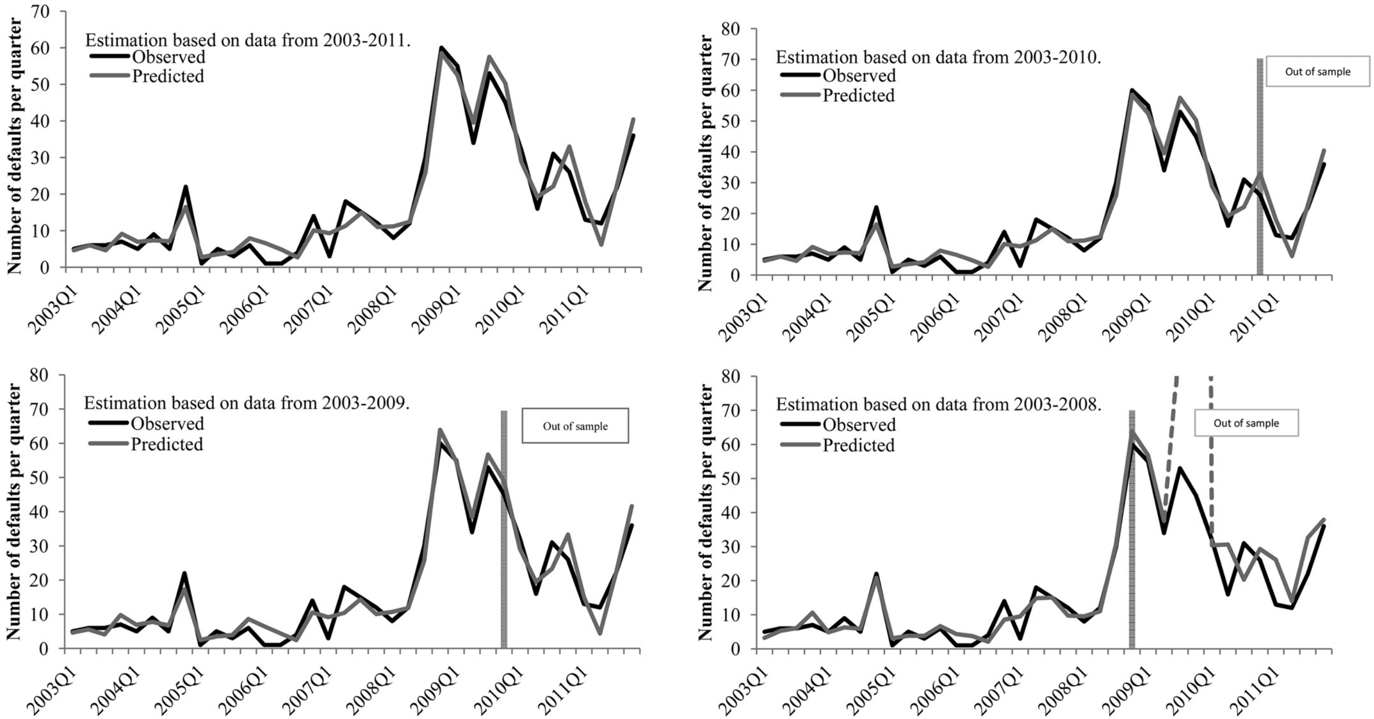
return of the Stoxx 50 index, and the European GDP growth rate appear with an opposite sign in the preferred model compared with a model where they enter separately.

Nonetheless, a positive return of the OMX stock market would, controlling for other macroeconomic effects, imply a lower number of default occurrences one year after. An increased spread between CIBOR and the policy rate would be associated with an increased number of default occurrences, thereby supporting the notion that the higher funding costs of the banks would generally be passed through to clients. Both growth in house prices and changes in the consumer confidence index tend to be negatively linked to defaults, illustrating the importance of demand-side effects. Capacity utilization is also negatively associated with default occurrences, meriting the interpretation that a higher level of idle capacity could result in price competition that would ultimately lead a number of firms to default.

Figure 4 illustrates the relationship between the observed and predicted number of defaults, taking into account both the firm-specific variables and the macroeconomic variables in model 7 of table 5. Adding macroeconomic variables as explanatory factors improves the model's ability to predict the cyclical variation in quarterly default occurrences. To rule out that the good fit of the preferred model is merely a result of overfitting, the out-of-sample prediction obtained from estimating the same model on only part of the data supports the chosen model. Panels B and C of the figure estimate the model on the sample excluding observations from 2010 and both 2010 and 2011, respectively. The obtained coefficients from the models estimated on the reduced samples are then used to estimate the aggregate intensities for all thirty-six quarters, thereby generating out-of-sample predictions. Hence, panel B of the figure shows one year and panel C shows two years of out-of-sample prediction. The out-of-sample prediction based on the reduced-sample estimation adequately captures both the level and cyclical variation in default rates.

Panel D of figure 4 shows prediction based on excluding the years 2009, 2010, and 2011 from the estimation. For the out-of-sample prediction in panel D, large deviations occur in 2009 (which pertains to 2008 covariates observations because of the one-year lag). However, it should be emphasized that the latter model has been fitted to a

Figure 4. Default Prediction Based on the Preferred Cox Model with Both Accounting Ratios and Macroeconomic Variables



Notes: All panels show the observed quarterly number of defaults in the sample along with a predicted number of defaults based on the preferred Cox model with both firm-specific and macroeconomic covariates (model 7 in table 5). In the top left panel, the model is estimated using the full sample from 2003 to 2011. In the other panels the model is estimated on shorter subsamples, allowing in each case out-of-sample prediction on the remaining part of the full sample.

period of economic expansion, and therefore it is of little surprise that the model cannot be used to predict future defaults in a period of economic contraction. This finding also highlights the importance of estimating default-predicting occurrences on a full business cycle.

6.3 Ranking Firms with Respect to Default Likelihood

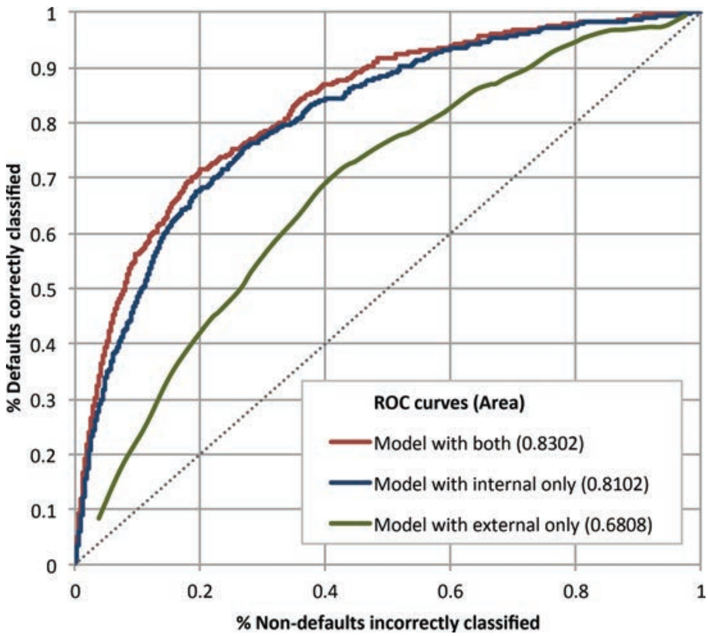
The out-of-sample estimation results presented in figure 4 showed that the preferred Cox model including both firm-specific and macroeconomic variables adequately captures the level and cyclicality of defaults. A common way of illustrating the predictive power of different models based on the same data is to plot the receiver operating curve (ROC), as shown in figure 5. The curve illustrates the percentage of defaults that are correctly classified as defaults on the vertical axis against the percentage of non-defaults that are mistakenly classified as defaults on the horizontal axis for all possible cutoff points. The area under the curve (AUC) is then used as a measure of the model goodness of fit, where a value of 1.0 implies a model with perfect discriminatory ability and a value of 0.5 is a completely random model.

In terms of discriminatory power, the addition of macroeconomic variables does not improve the model's ability to effectively determine which firms eventually default beyond what is implied by the accounting ratios. From the ROC curves, we see that the model with both macroeconomic and firm-specific variables is only marginally better in correctly determining defaults compared with the model with just firm-specific variables. This confirms that it is the firm-specific characteristics that provide the ordinal ranking of firms and therefore also ultimately determine *which* firms actually default. Including the macroeconomic factors only improves the model's ability to capture the cyclicality in the aggregate default rate, which is related to *when* defaults occur.

6.4 Allowing for Time-Varying Effects of Firm-Specific Variables

We now use the additive Aalen model to judge whether allowing for time-varying effects of firm-specific variables changes our

Figure 5. Comparison of Firm-Ranking Accuracy for Different Covariate Specifications

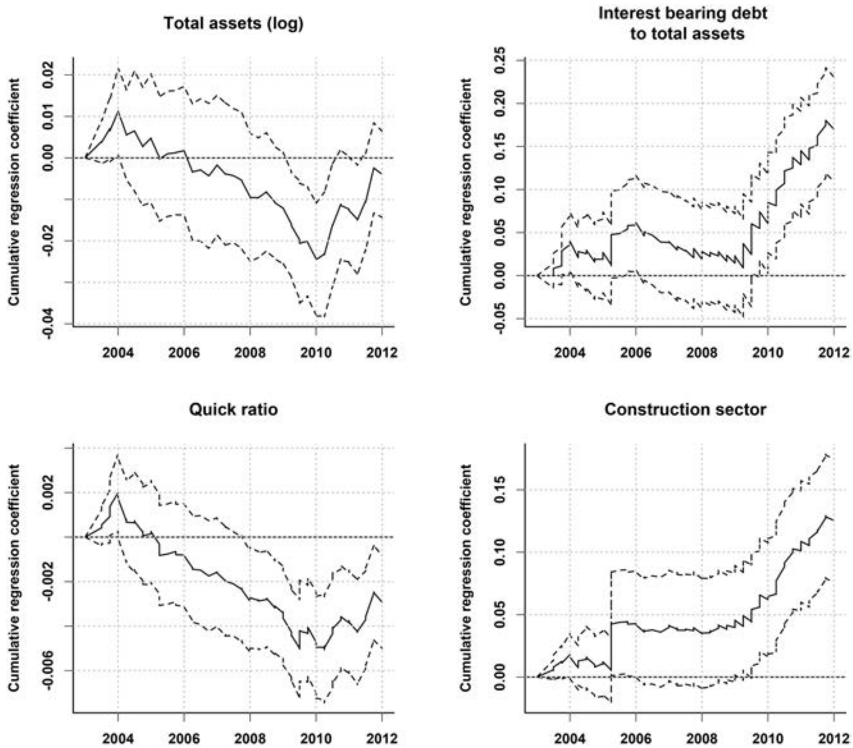


Notes: The figure shows receiver operating characteristic (ROC) curves for Cox models with different covariate specifications fitted to the sample. Each curve illustrates the model's ability to correctly discriminate between defaults and non-defaults, and plots the percentage of correctly classified defaults (true positives) against the percentage of incorrectly classified non-defaults (false positives) at all possible cutoff points of default intensity. The area under each curve serves as a goodness-of-fit measure, where a value of 1 means a model with perfect discriminatory ability, while a value of 0.5 means a model that discriminates based on a random guess.

conclusions regarding which firm-specific and macroeconomic variables significantly affect default intensities.

We initially fit an additive model for our entire sample of firms, including the same covariate specification as our final Cox model (model 7 of table 5), and with time-varying coefficients for the firm-specific covariate. The null hypothesis of a time-constant marginal effect is rejected at standard significance levels for the following four firm-specific variables: firm size (log of total assets), interest-bearing debt to total assets, the quick ratio, and the construction-sector

Figure 6. Cumulative Regression Coefficients from Aalen Analysis of Full Sample



Notes: The panels show (cumulative) estimation results for the significantly time-varying firm-specific covariates from an analysis based on the additive Aalen model including the same covariate specification as our final Cox model (model 7 of table 5). All variables are lagged one year to allow for one-year prediction. The dotted lines are asymptotic 95 percent pointwise confidence bands.

indicator. The time-varying marginal effects of these variables are shown as cumulative regression coefficients with 95 percent pointwise confidence bands in figure 6. When interpreting these effects, one should focus on the *slopes* of the cumulative coefficients, which estimate the regression coefficients themselves. The plots show that smaller firms, firms with higher interest-bearing debt to total assets, firms with a lower quick ratio, and firms within the construction sector have higher default probabilities, and particularly so during the financial crisis of 2008–10. Importantly, firm size has a mostly

Table 6. Constant Regression Coefficients from Aalen Analysis of Full Sample

Variables (All Lagged One Year)	Coef.	Se.
Intercept	0.0025	0.0407
Years Active in the Bank	-2.05×10^{-5}***	5.34×10^{-6}
Net Income to Total Assets	−0.0140***	0.0017
PPE to Total Assets	−0.0014**	0.0006
Dummy for Negative Equity	0.0078***	0.0012
Wholesale and Retail Trade, Dummy	0.0005	0.0004
Manufacturing, Dummy	0.0012**	0.0004
OMX Stock Market Return	−0.0036**	0.0016
OMX Stock Market Volatility	−0.0064	0.0080
Difference CIBO – Policy Rate, Pct. Point	0.0059	0.0510
Yield-Curve Slope 10y – 3m, Pct. Point	0.0110	0.0111
Growth in House Prices	−0.0006	0.0024
Change in Consumer Confidence Indicator	−0.0105**	0.0047
Change in Capacity Utilization	−0.0077	0.0085
Stoxx50 Stock Market Return	0.0042***	0.0012
EU-27 Real GDP Growth	0.0050	0.0077

Notes: The table shows the estimation results for the time-constant regression coefficients for the firm-specific and macroeconomic variables from an analysis based on the additive Aalen model including the same covariate specification as our final Cox model (model 7 of table 5). All variables are lagged one year to allow for one-year prediction. *, **, and *** denote parameter significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Parameter significance is based on robust standard errors.

negligible marginal effect for most of our sample period (consistent with the results from our analysis based on the Cox model), except for around the financial crisis of 2008–10, where the effect becomes significantly negative. Hence, contrary to the assumption in the Basel Accords, we find that larger firms were safer, not riskier, during the most recent recession.

The estimation results for the time-constant regression coefficients from the additive model fitted to the entire sample are given in table 6. We see that all coefficients corresponding to firm-specific variables have the same sign as in our final Cox model (model 7 of table 5) and roughly the same significance level. The macroeconomic variables, however, appear to have lost much of their importance compared with the analysis based on the Cox models. In the additive setting, only the OMX stock market return, the change in the consumer confidence indicator, and the Stoxx50 stock market return have significant marginal effects. The latter is of the reversed sign compared with intuition but is nonetheless consistent with results for public firms found by Duffie, Saita, and Wang (2007), Duffie et al. (2009), Lando and Nielsen (2010), and Figlewski, Friedman, and Liang (2012), amongst others.[6] Importantly, we see that allowing for time-varying effects for firm-specific variables leaves little room for additional explanatory power for macroeconomic variables. This is a consequence of the fact that the time variation in the effects shown in figure 6 is highly correlated with indicators of macroeconomic conditions, leaving little room for additional explanatory power of macroeconomic variables.

In sum, the analysis based on the additive model suggests that allowing for time variation in the effects of firm-specific variables is, indeed, an important component of the cyclicality of default probabilities that is missing from our preferred Cox model.

7. The Macroeconomy's Impact on Small and Large Firms' Default Risk

The results of the previous section show that our final Cox model specification including both firm-specific and macroeconomic variables is able to both accurately rank firms and predict the aggregate default rate over time. In this section, we first use this Cox model to investigate whether our data supports the assumption underlying the Basel II and III Accords that the default probabilities of smaller firms are less affected by the macroeconomy. We then redo

[6]Giesecke, Lando, and Medhat (2013) show that univariately significant but multivariately insignificant or even reversed effects may be observed for macroeconomic variables if these have indirect effects mediated through other covariates included in the models. This is, in particular, the case for stock market returns.

our analysis using the Aalen model, in order to judge whether allowing for time-varying effects of firm-specific variables changes our conclusions.

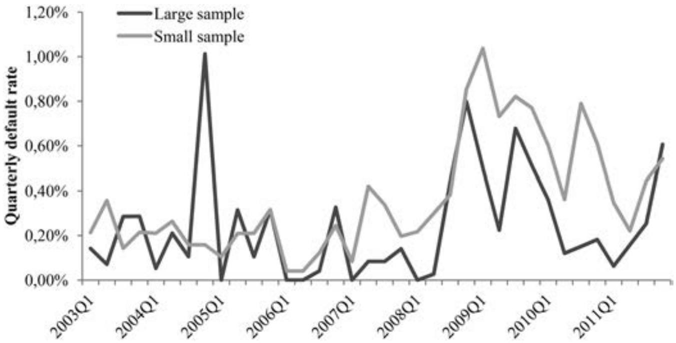
Our tests examine the extent to which the default probabilities of small and large firms are affected differently by macroeconomic variables. To do so, we choose to split the sample into two subsamples, corresponding to "small" and "large" firms. Based on the year that a particular firm enters the sample, it is classified as "small" ("large") if its first-year asset level is below (above) the median asset level that year. We use assets instead of revenue as our measure of size because, as discussed in section 2.1, Danish reporting standards allow firms below a certain size to refrain from reporting revenue. Note, however, that our use of asset level as the measure of size is in line with article 274 of the Basel Committee on Banking Supervision's (2006) report.[7]

Our choice to divide the sample based on the median asset level at a firm's entry year is done to ensure approximately equal sample sizes with a sufficient amount of default events in each subsample, as well as to allow for the classification of firms as "small" or "large" in a predictable manner. However, we later show (in the section on robustness and model checks) that our results are robust to a division into four subsamples based on asset value quartiles for the entry year.

Figure 7 shows the aggregate quarterly default rate in each of the subsamples of small and large firms. The two default rates are seen to exhibit a high degree of co-movement throughout our entire sample period. In the beginning of our sample period, the two default rates are seen to be virtually indistinguishable except for the spike in the default of large firms around 2005. However, near the end of our sample period, the default rate of small firms tends to be systematically higher than the default rate of large firms, indicating that small firms were hit the hardest by the financial crises of 2008–10.

[7]The article's wording is as follows: "Subject to national discretion, supervisors may allow banks, as a failsafe, to substitute total assets of the consolidated group for total sales in calculating the SME threshold and the firm-size adjustment. However, total assets should be used only when total sales are not a meaningful indicator of firm size."

Figure 7. Aggregate Quarterly Default Rate in Each of the Subsamples of Small and Large Firms



Notes: This figure shows the aggregate quarterly default rate in each of the subsamples of small and large firms. Based on the year a firm enters the sample, it is classified as "large" if its first-year asset level is above the median asset level that year. Similarly, a firm is classified as "small" if its asset value at the time of entry is below the median asset level that year.

7.1 Macro-sensitivity Analysis Based on the Cox Model

Table 7 shows estimation results for the final Cox model specification including both firm-specific and macroeconomic variables fitted to each of the two subsamples. With the exception of OMX stock market volatility, the magnitude of all macroeconomic effects is larger for large firms than for small firms. The difference between the coefficients of macroeconomic factors for the two subsamples is significant for the slope of the yield curve, growth in house prices, and European GDP growth, and all are larger in magnitude in the subsample of large firms.[8] Hence, if we suppose there exists a large and a small firm whose only difference is their size (which is in principle possible, since all accounting ratios in our models are relative to total assets), the apparent interpretation of these results is that the small firm's default intensity is less exposed to macroeconomic fluctuations.

[8] In an unreported analysis, we find that the importance of the European GDP growth for the large firms is merited by the tendency of large firms in the sample to engage more actively in exports.

Table 7. Estimation Results for Cox Models Fitted to the Subsamples of Small and Large Firms

	Small Firms		Large Firms		Comparison				
Variables (All Lagged One Year)	**Coef.**	**Se.**	**Coef.**	**Se.**	**Same Sign**	**Magnitude**	**Sig. Diff.**	**PEA**	**APE**
Intercept	−8.290***	0.676	−9.057***	0.765	Yes	Large	No	Small	Small
Years Active in the Bank	−0.008**	0.004	−0.012***	0.004	Yes	Large	No	Small	Small
Log of Total Assets	0.122**	0.056	−0.006	0.043	No	NA	No	NA	NA
Interest-Bearing Debt to Total Assets	1.048***	0.224	1.628***	0.351	Yes	Large	No	Small	Small
Net Income to Total Assets	−1.788***	0.302	−2.432***	0.422	Yes	Large	No	Small	Small
Quick Ratio	−0.112	0.074	−0.462***	0.137	Yes	Large	Yes	Large	Large
PPE to Total Assets	0.040	0.197	0.508**	0.255	Yes	Large	No	Large	Large
Dummy for Negative Equity	0.782***	0.199	0.269	0.291	Yes	Small	No	Small	Small
OMX Stock Market Return	−0.045***	0.015	−0.045**	0.018	Yes	Large	No	Small	Small
OMX Stock Market Volatility	−0.102***	0.038	−0.100**	0.049	Yes	Small	No	Small	Small
Difference CIBOR – Policy Rate, Pct. Point	1.655***	0.447	2.463***	0.553	Yes	Large	No	Small	Small
Yield-Curve Slope 10y – 3m, Pct. Point	0.372***	0.130	0.869***	0.154	Yes	Large	Yes	Small	Large
Growth in House Prices	−0.040	0.032	−0.200***	0.047	Yes	Large	Yes	Large	Large
Change in Consumer Confidence Indicator	−0.083***	0.026	−0.122***	0.036	Yes	Large	No	Small	Small
Change in Capacity Utilization	−0.194**	0.086	−0.420***	0.106	Yes	Large	No	Small	Large
Stoxx50 Stock Market Return	0.027***	0.009	0.043***	0.013	Yes	Large	No	Small	Small
EU-27 Real GDP Growth	0.101*	0.055	0.410***	0.070	Yes	Large	Yes	Large	Large

(*continued*)

Table 7. (Continued)

	Small Firms		Large Firms		Comparison				
Variables (All Lagged One Year)	**Coef.**	**Se.**	**Coef.**	**Se.**	**Same Sign**	**Magnitude**	**Sig. Diff.**	**PEA**	**APE**
PEA Scaling Factor	0.0023		0.0009						
APE Scaling Factor	0.0041		0.0025						
Number of Observations	91,182		101,014						
Number of Firms	5,333		5,338						
Number of Defaults	376		257						
Sector Effects	Yes		Yes						

Notes: The table shows parameter estimates, standard errors, model summary statistics, and comparison criteria for Cox models of the quarterly default intensity for small and large firms in the sample. The covariate list corresponds to the preferred model including both firm-specific and macroeconomic variables (model 7 in table 5). All variables are lagged one year to allow for one-year prediction. In each year, an entering firm is deemed "small" if its book value of assets is under the median assets level for that particular year—otherwise, it is deemed "large." *, **, and *** denote parameter significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Parameter significance is based on standard errors clustered at the firm level. "Same Sign" indicates whether or not the estimated coefficients are of the same sign for small and large firms; "Magnitude" indicates which type of firm has the largest estimated coefficient; and "Sig. Diff." indicates whether or not the coefficients for small and large firms are significantly different at the 5 percent level using a Welch t-test. Finally, "PEA" indicates whether the partial effect at the average is largest for the large firms or the small firms, while "APE" indicates whether the average partial effect is largest for the large firms or the small firms.

On the other hand, the estimation results for the two subsamples also show substantial differences with regards to the coefficients of the firm-specific variables: Firm size appears with a significant positive coefficient for small firms but an insignificant (yet negative) coefficient for large firms; neither the quick ratio nor the ratio of tangible assets to total assets has significant effects for small firms, whereas they have significant effects for large firms; and, finally, negative equity has a significant effect for small firms but not for large firms.

While the results for the macroeconomic variables may corroborate the lower asset correlation adopted in Basel II and III for SMEs, the direct comparison of coefficients in the two subsamples ignores the fact that a covariate's marginal effect in a non-linear model, such as a Cox regression, depends on the values of all the other covariates. This implies that comparing the coefficients for the macroeconomic variables in the two subsamples is potentially problematic, because such a comparison fails to take into account that the firm-specific characteristics for the small and large firms are generally different and have different effects on default intensity.

To elaborate, note that the marginal effect in a Cox regression of a change in the jth macroeconomic variable on the default intensity of firm i is given by

$$\frac{\partial \lambda(\mathbf{X}_{it}, \mathbf{Z}_t)}{\partial Z_{jt}} = \gamma_j\, Y_{it} \exp\left(\boldsymbol{\beta}^\top \mathbf{X}_{it} + \boldsymbol{\gamma}^\top \mathbf{Z}_t\right) = \gamma_j\, \lambda(\mathbf{X}_{it}, \mathbf{Z}_t),$$

which depends on all the characteristics of firm i through $\mathbf{X}_{it}$, as well as all other macroeconomic variables through the dependence on $\mathbf{Z}_t$.

A somewhat crude way to facilitate comparison between subsamples in non-linear models, such as a Cox regression, is to compute the marginal effect of a covariate at "average levels" in each of the subsamples. This gives rise to the *partial effect at the average* (PEA) and the *average partial effect* (APE)—see, for instance, Wooldrige (2009, 582–83). In the setting of an intensity model, the PEA plugs a subsample's average covariate values into the subsample's estimated intensity, while the APE takes the average across the estimated intensity values for each subsample. Due to the non-linearity of the intensity, Jensen's inequality implies that the two ways of averaging will generally produce different results.

In our analysis, the PEA is a measure of a covariate's marginal effect for the "average firm" and at "average macroeconomic levels" in each of the two subsamples of small and large firms. We thus compute the PEA for the jth macroeconomic variable in subsample k as

$$\mathrm{PEA}_{kj} = \gamma_{kj} \underbrace{\exp\left(\widehat{\boldsymbol{\beta}}_k \overline{\mathbf{X}}_k + \widehat{\boldsymbol{\gamma}}_k^{\top} \overline{\mathbf{Z}}_k\right)}_{=s_k^{\mathrm{PEA}}},$$

where $k \in \{\text{small}, \text{large}\}$, $\overline{\mathbf{X}}_k$ is the average firm-specific covariate vector for firms in subsample k, $\overline{\mathbf{Z}}_k$ is the average macroeconomic covariate vector in subsample k, $\widehat{\boldsymbol{\beta}}_k$ and $\widehat{\boldsymbol{\gamma}}_k$ are the estimated regression coefficients in subsample k, and s_k^{PEA} is a subsample-specific scaling factor for each PEA. On the other hand, the APE is a measure of a covariate's marginal effect at the "average intensity level" across firms and time in each of the two subsamples. The APE for the jth macroeconomic variable in subsample k is thus computed as

$$\mathrm{APE}_{kj} = \gamma_{kj} \underbrace{\frac{1}{T} \int_0^T \frac{1}{|k(t)|} \sum_{i \in k(t)} \exp\left(\widehat{\boldsymbol{\beta}}_k^{\top} \mathbf{X}_{it} + \widehat{\boldsymbol{\gamma}}_k^{\top} \mathbf{Z}_t\right) dt}_{=s_k^{\mathrm{APE}}},$$

where $k(t)$ denotes the firms belonging to subsample $k \in \{\text{small}, \text{large}\}$ at time t, and s_k^{APE} is again a subsample-specific scaling factor for each APE.

The two rightmost columns of table 7 show the PEAs and APEs for each covariate in each of the two subsamples. Focusing on the effects of the macroeconomic variables, the PEA suggests that most macro effects are, on average, stronger in the sample of small firms, while the APEs suggest that it is entirely dependent on the macrovariable at hand whether its average effect is stronger for small or large firms.

In sum, while the direct comparison of regression coefficients indicates that smaller firms are less cyclical than larger firms, the more refined analysis based on the PEA and APE, which takes the nonlinearity of the Cox model into account, indicates that small firms may "on average" be as cyclical, or perhaps even more cyclical, than large firms. To rule out that this conclusion somehow hinges on the Cox specification, we now investigate macro-sensitivity of default intensities using an additive intensity regression model.

7.2 Macro-sensitivity Analysis Based on the Additive Aalen Model

Due to the linearity of its effects, the additive model allows us to directly compare the effects of macroeconomic variables for small and large firms without the need for averaging techniques.

Recall that our specification of the additive model for the default intensity of firm i is given by

$$\lambda(\mathbf{X}_{it}, \mathbf{Z}_t) = \boldsymbol{\beta}(t)^\top \mathbf{X}_{it} + \boldsymbol{\gamma}^\top \mathbf{Z}_t.$$

It follows that the marginal effect of the jth firm-specific variable is given by

$$\frac{\partial \lambda(\mathbf{X}_{it}, \mathbf{Z}_t)}{\partial X_{ij,t}} = \beta_j(t),$$

which is time varying but otherwise independent of other variables and their effects. Similarly, the marginal effect of the jth macroeconomic variable is

$$\frac{\partial \lambda(\mathbf{X}_{it}, \mathbf{Z}_t)}{\partial Z_{j,t}} = \gamma_j,$$

which is time constant but, again, independent of other variables and their effects. Hence, in contrast to the multiplicative effects in the Cox model, the linear effects in the additive model are easy to interpret and compare across subsamples without the need for averaging techniques. We will therefore use the additive model to check the assumption that the default probabilities of smaller firms are less sensitive to macroeconomic cycles.

We now redo our macro-sensitivity analysis by fitting an additive model including the preferred firm-specific and macroeconomic variables to each of the two subsamples of small and large firms. The definition of small and large firms is the same as for the subsample analysis based on the Cox model in section 4.1. Even though the macroeconomic variables did not have much explanatory power in the additive model fitted to the entire sample, one could still imagine that some macroeconomic variables had significant additive effects in the subsample of large firms but not in the subsample of small firms. This would be evidence that macro-dependence differs for small and large firms, as is the working assumption in the Basel Accords.

Estimating the additive model in each of the two subsamples did not change the conclusions regarding which firm-specific variables have significantly time-varying effects. Furthermore, the time-varying effects within the two subsamples were virtually indistinguishable from their counterparts for the full sample shown in figure 6, and are therefore omitted here. In particular, in both subsamples, firm size has a mostly negligible marginal effect except during the financial crisis of 2008–10, where larger firms (within each subsample) have significantly lower default probabilities.

With regards to the time-constant effects, table 8 shows the estimation results for the time-constant regression coefficients in the subsamples of small and large firms. As previously mentioned, the linearity of the additive model allows us to directly compare coefficients across subsamples without the need for averaging techniques. Doing so, we see no difference in sign and virtually no difference in significance or magnitude in the two subsamples, and this is regardless of whether we consider the firm-specific variables or the macroeconomic variables.

In conclusion, our subsample analysis based on the additive model shows no evidence that the default probabilities of smaller firms are less sensitive to macroeconomic cycles. Our findings based on the additive model thus generally reiterate the ones based on the Cox model, but they do so in a framework that allows for cyclicality through the effects of firm-specific variables as well as direct comparison of effects across subsamples.

Table 8. Constant Regression Coefficients from Aalen Analysis of Small and Large Firms

Variables (All Lagged One Year)	Small Firms Coef.	Small Firms Se.	Large Firms Coef.	Large Firms Se.
Intercept	0.0597	0.1883	0.0600	0.1918
Years Active in the Bank	-1.83×10^{-5}*	1.14×10^{-5}	-1.70×10^{-5}	1.18×10^{-5}
Net Income to Total Assets	−0.0160***	0.0025	−0.0168***	0.0025
PPE to Total Assets	−0.0021**	0.0010	−0.0021**	0.0010
Dummy for Negative Equity	0.0082***	0.0016	0.0085***	0.0016
Wholesale and Retail Trade, Dummy	0.0003	0.0006	0.0003	0.0006
Manufacturing, Dummy	0.0011	0.0007	0.0011	0.0007
OMX Stock Market Return	−0.0084	0.0066	−0.0135*	0.0069
OMX Stock Market Volatility	0.0192	0.0324	0.0089	0.0317
Difference CIBOR – Policy Rate, Pct. Point	−0.3409	0.2317	−0.3048	0.2416
Yield-Curve Slope 10y – 3m, Pct. Point	0.0003	0.0524	0.0134	0.0560
Growth in House Prices	0.0136	0.0105	0.0098	0.0115
Change in Consumer Confidence Indicator	−0.0324*	0.0191	−0.0283*	0.0196
Change in Capacity Utilization	−0.0625	0.0402	−0.0713*	0.0404
Stoxx50 Stock Market Return	0.0111**	0.0050	0.0146**	0.0054
EU-27 Real GDP Growth	0.0049	0.0351	0.0131	0.0370

Notes: The table shows the estimation results for the time-constant regression coefficients for the firm-specific and macroeconomic variables from an analysis based on the additive Aalen model for the two subsamples of small and large firms. All variables are lagged one year to allow for one-year prediction. *, **, and *** denote parameter significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Parameter significance is based on robust standard errors.

8. Robustness and Model Check

In this section, we perform robustness checks and examine model fits. Because our main analysis is based on the Cox model, we will focus on this model. However, the results presented here are in general very similar when done based on the additive model.

8.1 Alternative Criteria for Defining Small and Large Firms

In our analysis of small and large firms in table 7, the criterion defining the two subsamples of small and large firms was the median asset level of the year where a particular firm enters the sample. We now investigate alternative criteria for defining the subsamples of small and large firms and how they affect the effects of our variables within the two subsamples.

Table 9 shows estimation results for Cox models within subsamples of small and large firms, where the subsamples are defined using different, time-varying size criteria. All the models use the covariate list corresponding to our preferred Cox model including both firm-specific and macroeconomic variables, i.e., model 7 in table 5.

Models A1 and A2 show the estimation results when the full sample is divided into subsamples of small and large firms according to the following criterion: A particular firm in a particular year is classified as small (large) if the firm's book asset value is below (above) the 25th percentile of the full sample's book asset values. This classification is time varying in the sense that the same firm can be classified as small in some years but large in other years. Similarly, models A3 and A4 show the estimation results when a particular firm in a particular year is classified as small (large) if the firm's book asset value is below (above) the 50th percentile of the full sample's book asset values. Lastly, models A5 and A6 show the estimation results when a particular firm in a particular year is classified as small (large) if the firm's book asset value is below (above) the 75th percentile of the full sample's book asset values.

A comparison of the estimation results within the three subsamples of small firms—i.e., models A1, A3, and A5—shows that the signs, magnitudes, and significance levels of the effects of both the firm-specific and the macroeconomic variables remain largely unaltered as the size criterion defining the subsample of small firms

Table 9. Estimation Results for Cox Models Fitted to Subsamples of Small and Large Firms Using Different, Time-Varying Size Criteria

Dependent Variable: Default (0/1)	**Cutoff: 25th Percentile**				**50th Percentile**				**75th Percentile**			
Variables (All Lagged One Year)	**Model A1**	**Se.**	**Model A2**	**Se.**	**Model A3**	**Se.**	**Model A4**	**Se**	**Model A5**	**Se.**	**Model A6**	**Se.**
Intercept	−9.00***	1.08	−8.62***	0.63	−8.32***	0.71	−8.58***	0.81	−8.14***	0.56	−5.80***	0.56
Years Active in the Bank	−0.01	0.01	−0.01***	0.00	−0.01**	0.00	−0.01**	0.00	−0.01***	0.00	0.00	0.00
Log of Total Assets	0.28**	0.11	−0.03	0.04	0.12**	0.06	−0.06	0.05	0.06	0.04	−0.29***	0.04
Interest-Bearing Debt to Total Assets	0.87***	0.26	1.67***	0.27	0.89***	0.22	2.15***	0.36	1.09***	0.21	2.25***	0.21
Net Income to Total Assets	−1.65***	0.34	−2.54***	0.34	−2.02***	0.28	−1.98***	0.51	−1.96***	0.26	−1.96**	0.26
Quick Ratio	−0.09	0.07	−0.36**	0.15	−0.12*	0.07	−0.53***	0.14	−0.16*	0.09	−0.55**	0.09
PPE to Total Assets	−0.01	0.26	0.29	0.20	0.23	0.19	0.11	0.25	0.31*	0.17	−0.18	0.17
Dummy for Negative Equity	0.93***	0.23	0.29	0.22	0.79***	0.19	0.19	0.31	0.67***	0.17	0.24	0.17
Construction	0.41**	0.20	1.34***	0.19	0.65***	0.17	1.56***	0.24	0.79***	0.15	1.81***	0.15
Wholesale and Retail Trade	0.01	0.19	0.42**	0.17	0.19	0.15	0.34	0.21	0.24*	0.14	0.30	0.14
Manufacturing	0.17	0.23	0.61***	0.17	0.29*	0.18	0.65***	0.21	0.42***	0.15	0.50	0.15
OMX Stock Market Return	−0.05***	0.02	−0.04***	0.02	−0.04***	0.01	−0.05**	0.02	−0.05***	0.01	−0.04	0.01
OMX Stock Market Volatility	−0.12**	0.05	−0.09**	0.04	−0.10***	0.04	−0.11**	0.05	−0.11***	0.03	−0.09	0.03
Difference CIBOR – Policy Rate, Pct. Point	1.45***	0.54	2.46***	0.44	1.79***	0.44	2.32***	0.55	1.94***	0.38	2.10***	0.38
Yield-Curve Slope 10y – 3m, Pct. Point	0.21	0.16	0.87***	0.13	0.36***	0.13	0.96***	0.16	0.47***	0.11	1.08***	0.11
Growth in House Prices	−0.03	0.04	−0.16***	0.04	−0.03	0.03	−0.25***	0.05	−0.06**	0.03	−0.33***	0.03
Change in Consumer Confidence Indicator	−0.08**	0.03	−0.11***	0.03	−0.09***	0.03	−0.11***	0.04	−0.10***	0.02	−0.10*	0.02
Change in Capacity Utilization in the Industrial Sector	−0.14	0.11	−0.41***	0.09	−0.16*	0.08	−0.50***	0.11	−0.24***	0.07	−0.48***	0.07
Stoxx50 Stock Market Return	0.04***	0.01	0.03***	0.01	0.03***	0.01	0.04***	0.01	0.03***	0.01	0.04*	0.01
EU-27 Real GDP Growth	0.07	0.07	0.35***	0.06	0.10*	0.05	0.44***	0.07	0.16***	0.05	0.56***	0.05

(*continued*)

Table 9. (Continued)

Dependent Variable: Default (0/1)	Cutoff: 25th Percentile				50th Percentile				75th Percentile			
Variables (All Lagged One Year)	**Model A1**	**Se.**	**Model A2**	**Se.**	**Model A3**	**Se.**	**Model A4**	**Se**	**Model A5**	**Se.**	**Model A6**	**Se.**
Number of Observations	48,049		144,147		96,083		96,113		144,150		48,046	
Number of Firms	3,817		8,291		6,420		5,530		8,728		2,762	
Number of Events	242		391		396		237		523		110	
Sector Effects	Yes		Yes		Yes		Yes		Yes		Yes	
QIC	2,739.2		4,733.6		4,586		2,881.4		6,119.9		1,365.5	
QICu	2,736.3		4,729.9		4,582		2,880.1		6,115.3		1,363.6	

Notes: The table shows parameter estimates, standard errors, and model summary statistics for Cox models of the quarterly default intensity for small and large firms in the sample. The covariate list corresponds to the preferred model including both firm-specific and macroeconomic variables (model 7 in table 5). All variables are lagged one year to allow for one-year prediction. The firm-specific variables are kept in all models to focus on the added explanatory effect of macroeconomic variables. In models A1 and A2, a particular firm in a particular year is classified as small (large) if the firm's book asset value is below (above) the 25th percentile of the full sample's book asset values (which is DKK 9,553,000). In models A3 and A4, the cutoff is instead the 50th percentile of the full sample's book asset values (which is DKK 28,222,000). Finally, in models A5 and A6, the cutoff is the 75th percentile of the full sample's book asset values (which is DKK 100,723,000). This classification is time varying in the sense that the same firm can be classified as small in some years but large in other years. *, **, and *** denote parameter significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Parameter significance is based on standard errors clustered at the firm level.

increases. The only notable changes for the firm-specific variables are for the size variable itself, which becomes smaller and insignificant as the size criterion defining a small firm increases, and for the manufacturing-sector dummy, which becomes larger and significant. The only notable changes for the macroeconomic variables are for the slope of the yield curve and the change in capacity utilization, which gain magnitude and significance as the size criterion defining a small firm increases. Furthermore, a comparison of models A1, A3, and A5 with the model for small firms in our main analysis in table 7 shows that using a time-varying classification of small firms entails no material changes to the signs, magnitudes, or significance levels of the effects of our variables within the subsamples of small firms.

Similarly, a comparison of the estimation results within table 9's three subsamples of large firms—i.e., models A2, A4, and A6—shows that the signs, magnitudes, and significance levels of the effects of both the firm-specific and the macroeconomic variables remain largely unaltered as the size criterion defining the subsample of small firms increases. Finally, a comparison of models A2, A4, and A6 with the model for large firms in our main analysis in table 7 shows that using a time-varying classification of large firms entails no material changes to the signs, magnitudes, or significance levels of the effects of our variables within the subsamples of large firms.

8.2 Independent Censoring and Entry Pattern

Given that data is only available from 2003 onwards, the existing stock of firms entering the sample in 2003 may potentially be of better average quality than the firms entering at a later point in time. This bias would violate the assumption of independent censoring. To address this issue, estimation was done on a reduced sample that excludes firms entering the sample in 2003 (where a considerable part of these entries ties to the existing stock of the bank clients). The results (not presented here but available upon request) are that all estimated coefficients remain significant and of the same sign as the final model (model 7 in table 5). To address the concern that the very low number of entries in 2005 might have an impact on the results, the final model specification was reestimated using two samples: one that excludes entries from 2005 and another that excludes all entries up until 2006. The estimates from these model fits (not presented here but available upon request) are still all significant and of the same sign as the model estimated on the full sample.

8.3 Lag Length

We have throughout chosen to focus on a lag length of one year for the covariates employed in our intensity models. However, one may believe that, for macroeconomic variables, this is not the appropriate lag, as aggregate changes may take longer to affect firms. This may, for instance, be because firms operate with a cash buffer that allows them to operate though an extended period of time before a default is observed. To address this concern, we reestimated the preferred model with all macroeconomic variables lagged eight quarters instead of four. The results (not presented here but available upon request) showed that the macroeconomic variables are generally less able to explain defaults when lagged eight quarters, as indicated by the loss in significance for most of the coefficients.

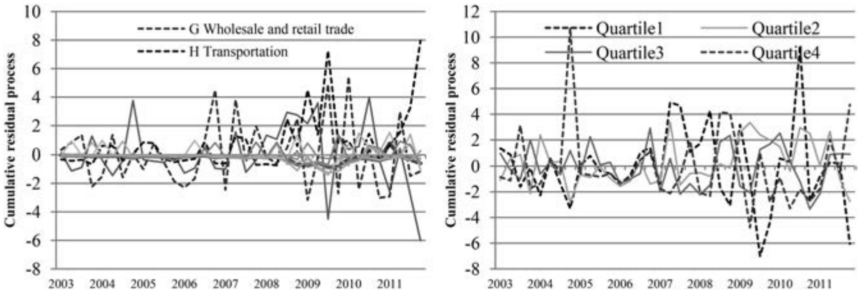
Still, this analysis does not consider the possibility that different macroeconomic variables are operating though different lag periods. Considering all possible combinations of lag periods would result in an extensive number of permutations of the model for us to check. Instead, we constructed a correlation matrix for the observed default rate and each of the macroeconomic time series lagged from zero to eight quarters. It generally shows that while the lag length of four quarters does not provide the highest correlation with the default rate for all macroeconomic variables, it appears that a unified lag period of four quarters is at least a very appropriate choice.

8.4 Grouped Martingale Residual Processes

We check the fit of our final model using the martingale residual processes (2). Specifically, we consider to what extent the model is systematically over- or under-estimating the default frequency in different sectors and size groups.

By definition, the martingale residual processes are the difference between the observed default frequency and the default frequency predicted by the model. Since a single firm can at most have one default event in our estimation setup, the firm-specific processes

Figure 8. Model Check Based on Grouped Martingale Residual Processes



Notes: The figure shows cumulative martingale residual processes for the preferred Cox model including both firm-specific and macroeconomic variables (model 7 in table 5). In the left panel, grouping corresponds to each firm's sector, which is stationary across the sample period, while in the right panel, the grouping is time varying and done according to asset quartiles.

contain too little information. However, when grouped in sufficiently large clusters, the increments of the grouped processes should not be systematically positive or negative if the model fit is adequate. An increasing grouped residual process would imply that the model is under-predicting the number of defaults for this particular group, whereas a decreasing grouped residual process would imply that the model is predicting too many defaults for this group.

The left panel of figure 8 shows grouped residual processes by sector as a function of time. We see that the residual processes fluctuate around zero with both positive and negative increments for all sectors. This is support for the model performing equally well for all sectors. However, noting that the sectors "wholesale and retail trade" and "transportation" have the largest deviances, we reestimated our final model excluding firms in these two sectors—the results (not presented here but available upon request) do not change.

The right panel of the figure shows grouped residual processes by asset quartiles as a function of time. We again observe no truly systematic deviations. We note, however, that the largest quarterly deviances occur in the largest and smallest quartiles, further motivating the point that default-prediction models should take firm size into account.

9. Conclusion

The Basel II and III Accords impose smaller capital charges on bank loans to SMEs by assuming a lower correlation in the formula for unexpected loss in a loan portfolio. The reduction in correlation for SMEs corresponds to an assumption that these firms have default probabilities that are less sensitive to the economy-wide factor. This paper investigates whether there is empirical support for this assumption. Using a default intensity regression framework, our results indicate that solely discriminating with respect to firm size, the default probabilities of small firms do in fact exhibit less sensitivity to macroeconomic cyclicality compared with the default probabilities of large firms, in the sense that the effects of macroeconomic variables are of smaller magnitude for smaller firms. However, when we account for differences in firm characteristics other than size, our results indicate that the default probability of the average small firm is as cyclical as or even more cyclical than the default probability of the average large firm. The results are robust to different regression models and different divisions of our sample into small and large firms. Our findings suggest that a reduction in capital charges based solely on firm size may imply a higher risk for banks with a high exposure to the SME segment.

References

Aalen, O. O. 1980. “A Model for Nonparametric Regression Analysis of Life Times.” In *Mathematical Statistics and Probability Theory*, Vol. 2 of Lecture Notes in Statistics, ed. J. R. W. Klonecki and A. Kozek, 1–25. New York: Springer-Verlag.

———. 1989. “A Linear Regression Model for the Analysis of Life Times.” *Statistics in Medicine* 8 (8): 907–25.

Aalen, O. O., Ø. Borgan, and H. K. Gjessing. 2008. *Survival and Event History Analysis: A Process Point of View.* 1st ed. New York: Springer-Verlag.

Altman, E. I. 1968. “Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy.” *Journal of Finance* 23 (4): 589–609.

Andersen, P. K., Ø. Borgan, R. D. Gill, and N. Keiding. 1992. *Statistical Models Based on Counting Processes.* Springer-Verlag.

Basel Committee on Banking Supervision. 2006. “International Convergence of Capital Measurement and Capital Standards — A Revised Framework (Comprehensive Version).” Technical Report, Bank for International Settlements (June).

Beaver, W. 1966. “Financial Ratios as Predictors of Failure.” *Journal of Accounting Research* (supplement titled *Empirical Research in Accounting: Selected Studies*) 4: 77–111.

Bharath, S. T., and T. Shumway. 2008. “Forecasting Default with the Merton Distance to Default Model.” *Review of Financial Studies* 21 (3): 1339–69.

Black, F., and M. Scholes. 1973. “The Pricing of Options and Corporate Liabilities.” *Journal of Political Economy* 81 (3): 637–54.

Bonfim, D. 2009. “Credit Risk Drivers: Evaluating the Contribution of Firm Level Information and of Macroeconomic Dynamics.” *Journal of Banking and Finance* 33 (2): 281–99.

Carling, K., T. Jacobson, J. Lind, and K. Roszbach. 2007. “Corporate Credit Risk Modeling and the Macroeconomy.” *Journal of Banking and Finance* 31 (3): 845–68.

Chava, S., and R. Jarrow. 2004. “Bankruptcy Prediction with Industry Effects.” *Review of Finance* 8 (4): 537–69.

Chava, S., C. Stefanescu, and S. Turnbull. 2011. “Modeling the Loss Distribution.” *Management Science* 57 (7): 1267–87.

Chionsini, G., J. Marcucci, and M. Quagliariello. 2010. “The Treatment of Small and Medium Enterprises in Basel 2: So Right, So Wrong?” Technical Report, Center for Economic Policy Research (March). Available at http://dev3.cepr.org/meets/wkcn/1/1737/papers/Marcucci.pdf.

Cox, D. R. 1972. “Regression Models and Life-Tables (with discussion).” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 34 (2): 187–220.

Dietsch, M., and J. Petey. 2004. “Should SME Exposures Be Treated as Retail or Corporate Exposures? A Comparative Analysis of Default Probabilities and Asset Correlations in French and German SMEs.” *Journal of Banking and Finance* 28 (4): 773–88.

Duffie, D., A. Eckner, G. Horel, and L. Saita. 2009. “Frailty Correlated Default.” *Journal of Finance* 64 (5): 2089–2123.

Duffie, D., and D. Lando. 2001. “Term Structures of Credit Spreads with Incomplete Accounting Information.” *Econometrica* 69: 633–65.

Duffie, D., L. Saita, and K. Wang. 2007. "Multi-period Corporate Default Prediction with Stochastic Covariates." *Journal of Financial Economics* 83 (3): 635–65.

European Commission. 2016. "Annual Report on European SMEs 2014/2015." Technical Report, Ref. Ares(2016)1791252 – 15/04/2016, European Commission.

European Parliament and the Council of the European Union. 2013. "Regulation (EU) No. 575/2013 of the European Parliament and of the Council of 26 June 2013 on Prudential Requirements for Credit Institutions and Investment Firms and Amending Regulation (EU) No. 648/2012." Technical Report. Available at http://eur-lex.europa.eu/legal-content/en/TXT/?uri=celex%3A32013R90575.

Figlewski, S., H. Frydman, and W. Liang. 2012. "Modeling the Effect of Macroeconomic Factors on Corporate Default and Credit Rating Transitions." *International Review of Economics and Finance* 21 (1): 87–105.

Giesecke, K., D. Lando, and M. Medhat. 2013. "The Macroeconomy and Credit Risk: Direct and Indirect Effects." Working Paper, Copenhagen Business School and Stanford University (September).

Jacobson, T., J. Lindé, and K. Roszbach. 2005. "Credit Risk versus Capital Requirements under Basel II: Are SME Loans and Retail Credit Really Different?" *Journal of Financial Services Research* 28 (1–3): 43–75.

Lando, D., M. Medhat, M. S. Nielsen, and S. F. Nielsen. 2013. "Additive Intensity Regression Models in Corporate Default Analysis." *Journal of Financial Econometrics* 11 (3): 443–85.

Lando, D., and M. S. Nielsen. 2010. "Correlation in Corporate Defaults: Contagion or Conditional Independence?" *Journal of Financial Intermediation* 19 (3): 335–72. Available at http://ssrn.com/paper=1338381.

Leland, H. 1994. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." *Journal of Finance* 49 (4): 1213–52.

Lopez, J. A. 2004. "The Empirical Relationship between Average Asset Correlation, Firm Probability of Default, and Asset Size." *Journal of Financial Intermediation* 13 (2): 265–83.

Martinussen, T., and T. H. Scheike. 2006. *Dynamic Regression Models for Survival Data.* 1st ed. New York: Springer-Verlag.

McDonald, C. G., and L. M. V. de Gucht. 1999. "High-Yield Bond Default and Call Risks." *Review of Economics and Statistics* 81 (3): 409–19.

Merton, R. C. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *Journal of Finance* 29 (2): 449–70.

Ohlsen, J. S. 1980. "Financial Ratios and the Probabilistic Prediction of Bankruptcy." *Journal of Accounting Research* 18 (1, Spring): 109–31.

Pesaran, M. H., T. Schuermann, B.-J. Treutler, and S. M. Weiner. 2006. "Macroeconomic Dynamics and Credit Risk: A Global Perspective." *Journal of Money, Credit and Banking* 38 (5): 1211–12.

Shumway, T. 2001. "Forecasting Bankruptcy More Accurately: A Simple Hazard Model." *Journal of Business* 74 (1): 101–24.

Wooldridge, J. M. 2009. *Introductory Economics—A Modern Approach.* 4th ed. South-Western.

Zmijewski, M. E. 1984. "Methodological Issues Related to the Estimation of Financial Distress Prediction Models." *Journal of Accounting Research* 22: 59–82.

Predicting Vulnerabilities in the EU Banking Sector: The Role of Global and Domestic Factors*

Markus Behn,[a] Carsten Detken,[a] Tuomas Peltonen,[b] and Willem Schudel[c]
[a]European Central Bank
[b]European Systemic Risk Board
[c]De Nederlandsche Bank

We estimate a multivariate early-warning model to assess the usefulness of private credit and other macrofinancial variables in predicting banking-sector vulnerabilities. Using data for twenty-three European countries, we find that global variables and in particular global credit growth are strong predictors of domestic vulnerabilities. Moreover, domestic credit variables also have high predictive power but should be complemented by other macrofinancial indicators such as house price growth and banking-sector capitalization that play a salient role in predicting vulnerabilities. Our findings can inform decisions on the activation of macroprudential policy measures and suggest that policymakers should take a broad approach in the analytical models that support risk identification and calibration of tools.

JEL Codes: G01, G21, G28.

*A previous version of the paper was titled "Setting Countercyclical Capital Buffers Based on Early-Warning Models—Would It Work?" (Behn et al. 2013). We would like to thank members of the ESRB Expert Group on Countercyclical Capital Buffers, as well as seminar participants at the European Central Bank, European Systemic Risk Board, the Banque de France Workshop on Countercyclical Capital Buffers on June 7, 2013 in Paris, and the 12th International Conference on Credit Risk Evaluation on September 26–27, 2013 in Venice for valuable comments and useful discussions. This paper is related to the comprehensive analysis conducted by the ESRB Expert Group on Countercyclical Capital Buffers. The views expressed are those of the authors and do not necessarily reflect those of the European Central Bank, the Eurosystem, De Nederlandsche Bank, or the European Systemic Risk Board. The usual disclaimer on errors applies here as well.
Author e-mails: markus.behn@ecb.europa.eu; carsten.detken@ecb.europa.eu; tuomas.peltonen@esrb.europa.eu; c.j.w.schudel@dnb.nl.

1. Introduction

Being faced with the longest and most severe financial crisis in decades, policymakers around the globe have actively searched for tools that could help to prevent or at least reduce the intensity of future financial crises. One such tool is the countercyclical capital buffer (CCyB), which aims to address cyclical systemic risk and is an integral part of the Basel III regulations and the EU Capital Requirements Directive (CRD IV). To measure the level of cyclical systemic risk, the Basel III framework promotes a methodology based on the ratio of aggregate credit to GDP (Basel Committee on Banking Supervision 2010), which has consequently featured prominently in policy decisions around the globe. However, while credit growth and the credit-to-GDP ratio are clearly important determinants of cyclical systemic risk, there are a number of other factors that can also indicate a buildup of vulnerabilities, which provides the motivation for this paper.

We assess the usefulness of credit and other macrofinancial variables for the prediction of banking-sector vulnerabilities in a multivariate framework, hence enabling a more informed decision on the activation of the CCyB. The Basel Committee on Banking Supervision (BCBS) guidelines are based on an analysis that uses a sample of twenty-six countries from all over the world, for which the credit-to-GDP gap (defined as the deviation of the credit-to-GDP ratio from its long-term trend) performs as the best single indicator in terms of signaling a coming financial crisis. However, the guidelines (or the underlying work by Drehmann, Borio, and Tsatsaronis 2011) rely on individual indicators and do not compare the predictive power of the credit-to-GDP gap to that of other potentially relevant variables related to risks to financial stability in a *multivariate framework*.[1] Acknowledging the potentially very large implications that this policy has for the international banking sector, our paper aims to fill this

[1]Several other potential shortcomings of the credit-to-GDP gap have been discussed in the literature. For example, Edge and Meisenzahl (2011) argue that gap measures are sensitive to the exact specification of the trending variable, in particular with regards to end-of-sample estimates of the credit-to-GDP ratio. For other critical views on the reliability or suitability of the credit-to-GDP gap in the context of the CCyB, see for example Repullo and Saurina (2011) and Seidler and Gersl (2011).

gap by estimating a multivariate early-warning model that includes private credit and other macrofinancial and banking-sector variables.

Our findings suggest that global variables and, in particular, global credit variables are strong predictors of macrofinancial vulnerabilities, providing good signals when used as single indicators and demonstrating consistent and significant effects in multivariate logit models. This concurs with the view that excessive global liquidity was one of the factors that contributed to the accumulation of financial vulnerabilities ahead of the global financial crisis (see, e.g., Committee on the Global Financial System 2011, International Monetary Fund 2013). The domestic credit-to-GDP gap also predicts vulnerabilities, although the effect is smaller than for the global credit variables. Despite the importance of credit variables, we also find evidence suggesting that other variables play a salient role in predicting vulnerable states of the economy. For example, domestic house price growth and global equity price growth are positively associated with macrofinancial vulnerabilities. Moreover, banking-sector variables exert significant effects: Strong banking-sector profitability may incur excessive risk-taking, leading to increased vulnerability, while a high banking-sector capitalization decreases the probability of entering a vulnerable state. This result is important for policymakers involved in setting the CCyB, as it reinforces the notion that higher bank capital ratios reduce the likelihood of financial vulnerability.

The results illustrate that even though credit variables are essential ingredients of early-warning models, other macrofinancial and banking-sector variables are important covariates to control for and to improve the predictive power of these models. Moreover, in an increasingly integrated economy, vulnerabilities that develop at a global level potentially transmit to countries around the world and should hence be taken into account when determining CCyB rates. The Basel III/CRD IV framework accounts for this by ensuring that the institution-specific CCyB rate is calculated as a weighted average of CCyB rates in countries to which the bank has exposures. Overall, policymakers should take a broad approach in the analytical models that support decisions on the activation of tools instead of focusing only on the domestic credit-to-GDP gap.

Our paper adds to the literature on early-warning models for financial and banking crises (see Alessi et al. 2015 and Holopainen and Sarlin 2016 for recent papers conducting horse races among existing methods). It builds on the so-called signaling approach, originally developed by Kaminsky, Lizondo, and Reinhart (1998), and extended by Demirgüç-Kunt and Detragiache (2000), Alessi and Detken (2011), Lo Duca and Peltonen (2013), and Sarlin (2013), and features a multivariate logit model to identify vulnerabilities in the banking system (see, e.g., Barrell et al. 2010, Karim et al. 2013, and Babecky et al. 2014). As a traditional early-warning model, our approach rests on the translation of predicted crisis probabilities from the logit model into binary signals that can then be evaluated given policymakers' preferences between type I (missing a crisis) and type II errors (false alarms of crises). Other recent contributions to the early-warning literature have focused on exploiting different methodologies to extract early warning models (Alessi and Detken 2014, Ferrari and Pirovano 2015), characterizations of the financial cycle (Drehmann, Borio, and Tsatsaronis 2013; Schüler, Hiebert, and Peltonen 2015; Galati et al. 2016), or machine learning approaches (Schüler, Hiebert, and Peltonen 2015, 2017; Holopainen and Sarlin 2016). While previous studies using multivariate early-warning models often focused on emerging markets (e.g., Frankel and Rose 1996; Demirgüç-Kunt and Detragiache 1998, 2000; and Manasse, Savona, and Vezzoli 2016) or a few large economies or individual countries (e.g., Hanschel and Monnin 2005, Ito et al. 2014, Castro, Estrada, and Martínez-Pagés 2016), our analysis is conducted for a sample of twenty-three European Union (EU) member states spanning the period from 1982 to 2012, hence informing possible decisions on CCyB rates in EU countries.

The remainder of the paper is organized as follows: We present our data set in section 2 and introduce the methodology in section 3. Estimation results and robustness checks are presented in section 4, while section 5 is reserved for our concluding remarks.

2. Data

This section introduces the data used for our study. We begin with the identification of vulnerable states, i.e., the dependent variable in the study, based on banking crises in the European Union. We then proceed by introducing the independent variables used in the empirical analysis. Finally, we present some descriptive statistics on the development of key variables around banking-sector crises in the sample countries.

2.1 Definition of Vulnerable States

The paper develops an early-warning model that attempts to predict vulnerable states of the economy from which—given a suitable trigger—banking crises could emerge. Thus, we are not trying to predict banking crises per se, even though we need to identify these crises in order to determine the vulnerable states. Specifically, we define a vulnerable state as the period twelve to seven quarters before the onset of a banking crisis. The time horizon accounts for the CCyB announcement period of twelve months that is specified in article 126(6) of the CRD IV, and for a time lag required to impose such a policy. At the same time, extending the horizon too far into the past may weaken the link between observed variation in the independent variables and the onset of banking crises. To analyze this, we provide a number of alternative time horizons in the robustness section.

In order to identify banking crises, we use the data set that has been compiled by Babecky et al. (2014) as part of a data collection exercise by the European System of Central Banks (ESCB) Heads of Research Group (labeled HoR database hereafter). This quarterly database contains information on banking crises in EU countries (except Croatia) between 1970:Q1 and 2012:Q4. The crisis index takes a value of 1 when a banking crisis occurred in a given quarter (and a value of 0 when no crisis occurred). The HoR database aggregates information on banking crises from "several influential papers," including (in alphabetical order): Caprio and Klingebiel (2003); Detragiache and Spilimbergo (2001); Kaminsky (2006); Kaminsky and Reinhart (1999); Laeven and Valencia (2008, 2010, 2012); Reinhart and Rogoff (2011, 2013); and Yeyati and Panizza (2011). The crisis indexes from these papers have subsequently been cross-checked with the ESCB Heads of Research before inclusion in the database. A list of the banking crisis dates for our sample countries based on this data set is provided in table 1. In the robustness section, we test the robustness of the results by regressing the benchmark model on banking crisis data provided by Reinhart and Rogoff (2011) and Laeven and Valencia (2012).

We set the dependent variable to 1 between (and including) seven to twelve quarters prior to a banking crisis as identified by the ESCB HoR database and to 0 for all other quarters in the data. In order to

Table 1. Data Availability and Crisis Dates

	Credit Variables	Other Variables	HoR Banking Crises
Austria	1982:Q1–2012:Q3	1986:Q4–2012:Q3	2008:Q4
Belgium	1982:Q2–2012:Q3	1982:Q1–2012:Q3	2008:Q3–2008:Q4
Czech Republic	1994:Q2–2012:Q2	—	1998:Q1–2002:Q2
Denmark	1982:Q2–2012:Q3	1992:Q2–2012:Q3	1987:Q1–1993:Q4, 2008:Q3–end of sample
Estonia	2005:Q1–2012:Q2	2005:Q2–2012:Q2	—
Finland	1982:Q2–2012:Q3	1987:Q2–2012:Q3	1991:Q1–1995:Q4
France	1982:Q2–2012:Q3	1992:Q2–2012:Q3	1994:Q1–1995:Q4, 2008:Q1–2009:Q4
Germany	1982:Q2–2012:Q2	1991:Q2–2011:Q4	2008:Q1–2008:Q4
Greece	2003:Q1–2012:Q2	2003:Q1–2012:Q2	2008:Q1–end of sample
Hungary	1997:Q1–2012:Q3	2002:Q1–2012:Q2	2008:Q3–2009:Q4
Ireland	1999:Q1–2012:Q3	1999:Q1–2010:Q4	2008:Q1–end of sample
Italy	1982:Q2–2012:Q3	1990:Q3–2012:Q2	1994:Q1–1995:Q4
Lithuania	2005:Q1–2012:Q2	2005:Q1–2012:Q2	2009:Q1–2009:Q4
Luxembourg	2004:Q2–2012:Q3	2004:Q2–2010:Q4	2008:Q2–end of sample
Malta	2006:Q2–2012:Q2	—	—
Netherlands	1982:Q2–2012:Q2	1982:Q1–2011:Q4	2008:Q1–2008:Q4
Poland	1997:Q1–2012:Q3	2003:Q1–2012:Q3	—
Portugal	1982:Q2–2011:Q4	1998:Q2–2011:Q4	—
Slovakia	2005:Q2–2012:Q2	—	—
Slovenia	2005:Q3–2012:Q2	—	—
Spain	1982:Q2–2012:Q3	1995:Q2–2012:Q3	1982:Q2–1985:Q3
Sweden	1982:Q2–2012:Q3	1986:Q2–2012:Q3	1990:Q3–1993:Q4, 2008:Q3–2008:Q4
United Kingdom	1982:Q2–2012:Q3	1988:Q2–2012:Q2	1991:Q1–1995:Q2, 2007:Q1–2007:Q4

Notes: The table shows the availability of credit and other variables as well as the crisis dates for the twenty-three countries in our sample. Credit variables are obtained from the BIS database for total credit to the private non-financial sector (see Dembiermont, Drehmann, and Muksakunratama 2013) and from Eurostat for those countries where the BIS data are not available. Other macrofinancial and banking-sector variables are obtained from various sources, including the BIS, IMF, and OECD. The crisis definitions are from the ESCB Heads of Research database described in Babecky et al. (2014).

overcome crisis and post-crisis bias (see, e.g., Bussière and Fratzscher 2006), we omit all country quarters that either witnessed a banking crisis or that fall within six quarters after a banking crisis.

2.2 Macrofinancial and Banking-Sector Variables

The panel data set used in the analysis contains quarterly macrofinancial and banking-sector data spanning 1982:Q2–2012:Q3 for twenty-three EU member states. The data is sourced through Haver Analytics and originally comes from the Bank for International Settlements (BIS), Eurostat, the International Monetary Fund (IMF), the European Central Bank (ECB), and the OECD (see the appendix for a detailed description). Table 1 provides an overview of the data availability for our main variables, while table 2 gives descriptive statistics for the variables included in our study.

Accelerated credit growth is one of the key variables in the Basel III/CRD IV framework, as it is seen as a sign of overheating that may be associated with systemic events in the banking sector (see also, e.g., Schularick and Taylor 2012, Drehmann and Juselius 2014, or López-Salido, Stein, and Zakrajšek 2016). To measure credit growth and levels, we follow Drehmann, Borio, and Tsatsaronis (2011) and use the long series on total credit and domestic bank credit to the private non-financial sector compiled by the BIS. This data includes borrowing from non-financial corporations, households, and non-profit institutions serving households. It aims to capture all sources of lending, independent of the country of origin or type of lender, and includes loans and debt securities such as bonds and securitized loans (see Dembiermont, Drehmann, and Muksakunratana 2013 for a description of the database). To our knowledge, the BIS credit series offers the broadest definition of credit provision to the private sector, while having been adjusted for data gaps and structural breaks. Our models account for credit growth and leverage, both at the domestic and at the global level. Credit growth is measured as a percentage (annual growth), while leverage is measured by the deviation of the credit-to-GDP ratio (using nominal GDP data) from its long-term backward-looking trend (see the appendix for details on the calculation). Global credit variables have been computed using a GDP-weighted average of the variable in question for several countries, including the United States, Japan,

Table 2. Descriptive Statistics

	Obs.	Mean	Std. Dev.	Min.	Max.
Dom. Credit Growth (qoq)	1,220	0.0228	0.0196	−0.0318	0.0989
Dom. Credit Growth (yoy)	1,220	0.0926	0.0662	−0.0690	0.3579
Dom. Credit Gap	1,220	0.1149	0.1186	−0.1570	0.4550
Dom Credit Growth (4q MA)	1,220	0.0232	0.0166	−0.0173	0.0897
Dom. Credit Growth (6q MA)	1,220	0.0232	0.0154	−0.0122	0.0813
Dom. Credit Growth (8q MA)	1,220	0.0233	0.0150	−0.0099	0.0805
Dom. Credit-to-GDP Ratio	1,220	1.2756	0.4259	0.4426	2.4829
Dom. Credit-to-GDP Gap	1,220	0.0346	0.0796	−0.1788	0.3249
Dom. Credit Growth – GDP Growth	1,220	0.0081	0.0171	−0.0508	0.0715
Glo. Credit Growth (qoq)	1,220	0.0152	0.0086	−0.0048	0.0335
Glo. Credit Growth (yoy)	1,220	0.0614	0.0289	−0.0113	0.1095
Glo. Credit Gap	1,220	0.0597	0.0431	−0.0101	0.1593
Glo. Credit Growth (4q MA)	1,220	0.0154	0.0071	−0.0028	0.0274
Glo. Credit Growth (6q MA)	1,220	0.0156	0.0069	−0.0021	0.0280
Glo. Credit Growth (8q MA)	1,220	0.0158	0.0065	0.0005	0.0274
Glo. Credit-to-GDP Ratio	1,220	0.7557	0.1193	0.5778	0.9933
Glo. Credit-to-GDP Gap	1,220	0.0158	0.0285	−0.0420	0.0676
Glo. Credit Growth – Glo. GDP Growth	1,220	0.0022	0.0235	−0.0492	0.0671
GDP Growth	919	0.0123	0.0088	−0.0232	0.0437
Inflation	919	0.0242	0.0166	−0.0108	0.1078
Equity Price Growth	919	0.0240	0.1199	−0.3759	0.3051
House Price Growth	919	0.0172	0.0289	−0.0735	0.1204
Banking-Sector Capitalization	756	0.0507	0.0161	0.0238	0.1088
Banking-Sector Profitability	756	0.0066	0.0040	−0.0142	0.0292
Gov. Bond Yield	862	0.0575	0.0237	0.0220	0.1385
Money Market Rate	862	0.0460	0.0292	0.0010	0.1643
Global GDP Growth	919	0.0117	0.0229	−0.0585	0.0616
Global Equity Price Growth	919	0.0135	0.0675	−0.3344	0.1122
Global House Price Growth	919	0.0066	0.0162	−0.0389	0.0539

Notes: The table shows descriptive statistics for the credit variables and the other macrofinancial indicators used in the empirical analysis. Credit variables are available for a longer period of time in most countries, which is why the number of observations is larger for them.

Canada, and all European countries that are in this study (see also Alessi and Detken 2011).

In order to test the importance of credit variables in a comparative fashion as well as to analyze the potential importance of other factors, we include a number of additional variables in our study. These variables are available for fewer observations than the credit variables, which is why the number of observations in the full model differs from the number of observations in models that include only credit variables (estimating credit models on the reduced sample yields results that are very similar to the ones for the full sample). To account for the macroeconomic environment and monetary stance, we include nominal GDP growth (domestic and global) and CPI inflation rates. Furthermore, we use data on annual equity and residential house price growth, both domestically and globally, to account for the common view that asset price booms can be associated with a buildup of vulnerabilities in the banking sector (see, e.g., Allen and Gale 2009, Brunnermeier and Oehmke 2013, or Jordà, Schularick, and Taylor 2015). Finally, to control for banking-sector profitability and solvency, we include aggregate bank capitalization (calculated by the ratio of equity over total assets) and aggregate banking-sector profitability (defined as net income before tax as a percentage of total assets).

As we are estimating binary choice models using panel data, non-stationarity of independent variables could be an issue (Park and Phillips 2000). We perform panel unit-root tests suggested by Im, Pesaran, and Shin (2003) as well as univariate unit-root tests developed by Dickey and Fuller (1979) in order to analyze the time-series properties of the variables of interest. In the panel unit-root test, the null hypothesis that all cross-sections contain unit roots can be rejected at least at the 10 percent level of confidence for all series except for the credit-to-GDP gap and global credit growth. We complement the panel unit-root analysis by using the Dickey and Fuller (1979) test country by country, and can reject the null hypothesis of a unit root for the credit-to-GDP gap at least at the 10 percent level for all countries except for Estonia, Lithuania, and Greece. Furthermore, in the country-by-country unit-root tests, we can reject the null hypothesis for the global credit-to-GDP gap at least at the 10 percent level for all countries except for Estonia and Lithuania, while for global credit growth the null hypothesis can be rejected for all countries. This implies that sample periods for individual countries seem to affect unit-root test results. Overall, the transformations done to the original variables, the results from the unit-root tests, and general economic theory make us confident that we have addressed potential non-stationarity concerns for the variables of interest.

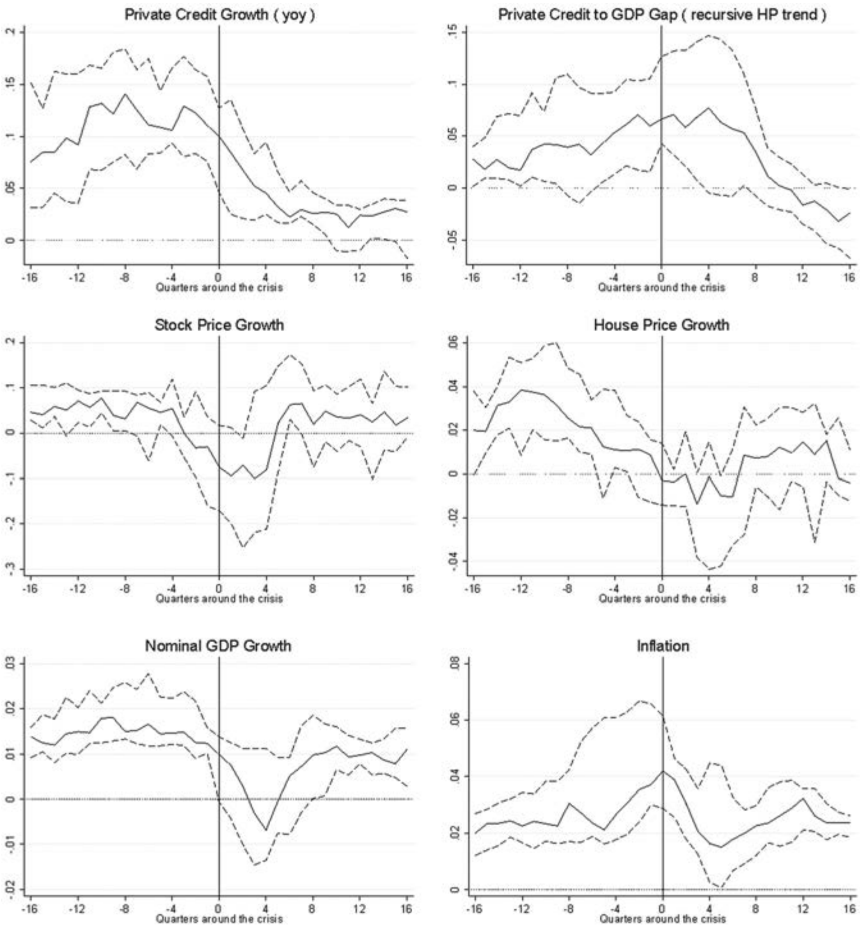
2.3 Development of the Key Variables

Before entering the discussion of the main results, we shortly present some descriptive statistics, which provide the context of our main argument of moving beyond credit variables when predicting macro-financial vulnerabilities. Figure 1 presents the average development of the six main variables of interest over time before and after the onset of a banking crisis. For the purpose of predicting crises, one would hope to find an indicator variable that (on average) peaks (or bottoms out, or at least changes direction) a number of quarters before a crisis, so that it can be used as a signal. In the current case of predicting a vulnerable state of the economy that precedes a potential banking crisis, we would be interested in variables that change direction a bit earlier before the onset of a crisis (i.e., two to three years before the crisis), so that policymakers can use this time to increase the resilience of banks.

In this context, we observe that among the six variables depicted here, the credit-to-GDP gap shows one of the least clear pictures in terms of signaling a coming crisis. On average, the credit gap increases slowly prior to a banking crisis and only starts falling about one year into the crisis. Yet, this does not need to be a very surprising development, as this variable is a ratio and therefore requires the numerator to grow more slowly (or decrease faster) than the denominator in order for the variable to decrease in value. The BCBS itself concedes that the credit-to-GDP trend may not capture turning points well (Basel Committee on Banking Supervision 2010). Consequently, the ratio will not fall unless credit falls faster than GDP, something that is not at all certain during banking crises. Still, it shows that purely from a descriptive perspective, any signal derived from the credit gap needs to come from the level of this variable (i.e., a threshold value), not from changes in its development.

Unlike the credit gap, credit growth (as depicted in percentage year-on-year growth) does appear to hit a peak about two years

Figure 1. Development of Key Variables around Banking Crises



Notes: The figure depicts the development of selected key variables around banking crises within the sample countries. The start date of a banking crisis is indicated by the vertical line, while the solid line shows the development in the median country and the dashed lines represent the countries at the 25th and 75th percentile, respectively.

before the onset of a banking crisis, even though its fall only becomes clear during the last pre-crisis year. A similar development can be observed in nominal GDP growth and equity price growth figures. These variables do peak before a crisis (on average), but the signal that a crisis is coming only becomes evident shortly before the crisis happens. This makes it difficult, at least from a descriptive point of view, to extract any strong signal from these variables. By some margin, residential house price growth outperforms the other domestic variables in terms of signaling "power" in this descriptive exercise. In our sample, the growth rate of residential house prices tends to peak about three years before a crisis happens on average, starting a clear descent (although prices are still rising) that lasts into the crisis where growth stalls. Based on this evidence, we would conclude that residential house prices would be a useful tool (at least much more useful than the other variables shown here) for decisions on the CCyB, as it clearly fulfills the early-warning requirement (one year of implementation plus one or two quarters of publication lag). So, at least from a descriptive standpoint, it is clear that it makes sense to gauge the developments of different macrofinancial variables to predict or signal coming crises. Whether this result holds in a more rigorous comparative (multivariate) framework will be discussed in the subsequent analysis.

3. Methodology

In this section we introduce the methodology used in the empirical analysis. We start by introducing the logistic regressions used in our multivariate framework. Thereafter, we explain how we evaluate individual indicators' and model predictions' usefulness for policymakers.

3.1 Multivariate Models

In order to assess the predictive abilities of credit, macrofinancial, and banking-sector variables in a multivariate framework, we estimate logistic regressions of the following form:

$$Prob(y_{it} = 1) = \frac{e^{\alpha_i + X'_{it}\beta}}{1 + e^{\alpha_i + X'_{it}\beta}}, \tag{1}$$

where $Prob(y_{it} = 1)$ denotes the probability that country i is in a vulnerable state, where a banking crisis could occur seven to twelve quarters ahead of quarter t. As independent variables, the vector

X_{it} includes credit and macrofinancial variables on the domestic and on the global level as well as domestic banking-sector variables (see section 2.2). The estimations also include country fixed effects, α_i, in order to account for unobserved heterogeneity at the country level.[2] Finally, we use robust standard errors clustered at the quarterly level in order to account for potential correlation in the error terms that might arise from the fact that global variables are identical across countries in a given quarter.[3]

The analysis is conducted as much as possible in a real-time fashion, meaning that only information that is available at a particular point in time is used. Therefore, all detrended variables have been calculated using backward-looking trends, and all explanatory variables have been lagged by one quarter, also to account for possible endogeneity. We are well aware that this simple procedure cannot crowd out all endogeneity-related bias, but we note that the dependent variable itself is an early-warning variable. The time horizon for which this variable is equal to 1 has been chosen in the context of our exercise and has not been exogenously determined. Therefore, we consider endogeneity to be a somewhat smaller problem in this study. Nevertheless, we have tested our models for different specifications of the dependent variable, both in terms of the pre-crisis period chosen (one to twelve/thirteen to twenty quarters before the onset of a crisis) and the definition and data source of banking crises in the robustness section.

3.2 Model Evaluation

Banking crises are (thankfully) rare events in the sense that most EU countries have encountered none or only one over the past two decades. Still, when they occur, banking crises tend to be very costly,

[2]We do not include time dummies for two reasons: First, only quarters where at least one country experiences a banking crisis could be used for identification in such a specification. As our sample includes many quarters where none of the countries experienced a crisis, the inclusion of time dummies would significantly reduce the sample size. Second, the focus in our paper is on the prediction of future banking crises. While time dummies might improve the ex post fit of a model, they are of little use for out-of-sample forecasting since they are not known ex ante.

[3]Clustering at the country level yields smaller standard errors, in particular for the global variables.

both directly through bailouts and fiscal interventions and indirectly through the loss of economic output that often tends to follow these crises (in particular, for systemic banking crises). Thus, policymakers have a clear incentive to be able to detect early enough potential signs of vulnerabilities that might precede banking crises in order to take measures to prevent further building up of vulnerabilities or to strengthen the resilience of the banking sector. Yet, at the same time, policymakers may not want to be signaling crises when in fact they do not happen afterwards. Doing so may (i) reduce the credibility of their signals, weakening decisionmaking and damaging their reputation, and (ii) needlessly impose costs on the banking sector, endangering credit supply. As a consequence, policymakers also have an incentive to avoid false alarms, i.e., they do not want to issue warnings when a crisis is not imminent. As pointed out by Alessi and Detken (2011), an evaluation framework for an early-warning model needs to take into account policymakers' relative aversion with respect to type I errors (not issuing a signal when a crisis is imminent) and type II errors (issuing a signal when no crisis is imminent).

The evaluation approach in this paper is based on the so-called signaling approach that was originally developed by Kaminsky, Lizondo, and Reinhart (1998) and extended by Demirgüç-Kunt and Detragiache (2000), Alessi and Detken (2011), Lo Duca and Peltonen (2013), and Sarlin (2013). In this framework, an indicator issues a warning signal whenever its value in a certain period exceeds a threshold τ, defined by a percentile of the indicator's country-specific distribution. Similarly, a multivariate probability model issues a warning signal whenever the predicted probability from this model exceeds a threshold $\tau \in [0, 1]$, again defined as a percentile of the country-specific distribution of predicted probabilities. In this way, individual variables and model predictions for each observation j are transformed into binary predictions P_j that are equal to 1 if the respective thresholds are exceeded for this observation and 0 otherwise. Predictive abilities can then be evaluated by comparing the signals issued by the respective variable or model with the actual outcome C_j for each observation. Each observation can be allocated to one of the quadrants in the contingency matrix depicted in figure 2: A period with a signal by a specific indicator can either be followed by a banking crisis seven to twelve quarters ahead (TP) or not (FP). Similarly, a period without a signal can be followed by a banking

Figure 2. Contingency Matrix

		Actual class C_{it}	
		1	0
Predicted class Q_{it}	1	True positive (TP)	False positive (FP)
	0	False negative (FN)	True negative (TN)

Notes: The figure shows the relationship between model prediction and actual outcomes. Observations are classified into those where the indicator issues a warning that is indeed followed by a banking crisis seven to twelve quarters ahead (TP), those where the indicator issues a warning that is not followed by a crisis (FP), those where the indicator issues no warning and there is no crisis seven to twelve quarters ahead (TN), and those where the indicator issues no warning although there is a crisis coming (FN).

crisis seven to twelve quarters ahead (FN) or not (TN). Importantly, the number of observations classified into each category depends on the threshold τ.

In order to obtain the optimal threshold τ, one needs to take the policymaker's preferences vis-à-vis type I errors (missing a crisis, $T_1(\tau) = FN/(TP + FN) \in [0, 1]$) and type II errors (issuing a false alarm, $T_2(\tau) = FP/(FP + TN) \in [0, 1]$) into account. This can be done by defining a loss function that depends on the two types of errors as well as the policymaker's relative preference for either type. The optimal threshold is then the one that minimizes the loss function. Taking into account the relative frequencies of crises $P_1 = P(C_j = 1)$ and tranquil periods $P_2 = P(C_j = 0)$, the loss function is defined as follows:

$$L(\mu, \tau) = \mu P_1 T_1(\tau) + (1 - \mu) P_2 T_2(\tau), \tag{2}$$

where $\mu \in [0, 1]$ denotes the policymakers' relative preference between type I and type II errors. A μ larger than 0.5 indicates that the policymaker is more averse to missing a crisis than to issuing a false alarm, which—in particular, following the recent financial crisis—is a realistic assumption in our view.

Using the loss function $L(\mu, \tau)$, the usefulness of a model can be defined in two ways. First, following the idea of Alessi and Detken

(2011) and as in Sarlin (2013), the absolute usefulness is defined as

$$U_a = min(\mu P_1, (1-\mu)P_2) - L(\mu, \tau). \quad (3)$$

Note that U_a computes the extent to which having the model is better than having no model. This is because a policymaker can always achieve a loss of $min(\mu P_1, (1-\mu)P_2)$ by either always issuing a signal (in which case $T_1(\tau) = 0$) or never issuing a signal (in which case $T_2(\tau) = 0$). The fact that P_1 is significantly smaller than P_2 in our sample (the share of observations that is followed by a banking crisis seven to twelve quarters ahead is approximately 10 percent) implies that, in order to achieve a high usefulness of the model, a policymaker needs to be more concerned about the detection of vulnerable states potentially preceding banking crises than the avoidance of false alarms. Otherwise, with a suboptimal performing model, it would easily pay off for the policymaker to never issue a signal given the distribution of vulnerable states and tranquil periods (see Sarlin 2013 for a detailed discussion of this issue).

A second measure, the relative usefulness U_r, is computed as follows (see Sarlin 2013):

$$U_r = \frac{U_a}{min(\mu P_1, (1-\mu)P_2)}. \quad (4)$$

The relative usefulness U_r reports U_a as a percentage of the usefulness that a policymaker would gain from a perfectly performing model.[4] The relative usefulness is our preferred performance indicator, as it allows the comparison of models for policymakers with different values for the preference parameter μ.[5]

4. Empirical Results

In this section we present the empirical results. We first explore the usefulness of credit variables for the identification of vulnerable

[4] A perfectly performing indicator would have $T_1 = T_2 = 0$, implying $L = 0$ and $U_a = min(\mu P_1, (1-\mu)P_2)$.

[5] We also employ receiver operating characteristics (ROC) curves and the area under the ROC curve (AUROC) for comparing performance of the early-warning models (see the appendix for details).

states of the banking sector, and proceed by extending the framework to a multivariate model including other macrofinancial and banking-sector indicators. Thereafter, we evaluate the out-of-sample performance of the estimated models and—finally—present some robustness checks.

4.1 Estimation and Evaluation

As the CRD IV regulations emphasize the role of credit variables for setting the countercyclical capital buffer rate—in particular, the role of credit growth and the credit-to-GDP gap—we start by evaluating the usefulness of these variables for the identification of vulnerable states within the EU banking sector.

4.1.1 Individual Indicators Based on Credit Growth and Credit-to-GDP Ratios

First, we evaluate the usefulness of domestic credit variables by using a simple signaling approach. Using a preference parameter of μ equal to 0.9, panel A of table 3 reports the optimal threshold for several credit variable indicators.[6] Given the optimal threshold, the table also shows the number of observations in each quadrant of the matrix depicted in figure 2; the percentage of type I and type II errors; and several performance measures, such as the absolute and the relative usefulness, the adjusted noise-to-signal (aNtS) ratio,[7] the percentage of vulnerable states correctly predicted by the indicator (% Predicted), the probability of a vulnerable state conditional

[6]A preference parameter of μ equal to 0.9 indicates a strong preference for the detection of crises by the policymaker. In our view this is a reasonable assumption, as the current crisis illustrated once more that financial crises often translate into large costs for the economy. As Sarlin (2013) points out, using a μ equal to 0.9 and simultaneously taking into account the unconditional probability of a crisis (which is about 10 percent in our sample) is equivalent to using a μ equal to 0.5 without adjusting for the unconditional probabilities (as in Alessi and Detken 2011 or Lo Duca and Peltonen 2013). Results for different values of μ are available upon request.

[7]The aNtS ratio is the ratio of false signals measured as a proportion of quarters where false signals could have been issued to good signals as a proportion of quarters where good signals could have been issued, or $(FP/(FP + TN))/(TP/(TP + FN))$. A lower aNtS ratio indicates better predictive abilities of the model.

Table 3. Evaluation of Individual Indicators

	μ	Threshold	TP	FP	TN	FN	T_1	T_2	Absolute Usefulness	Relative Usefulness	aNtS Ratio	% Predicted	Cond. Prob.	Diff. Prob.
A. Domestic Variables														
Dom. Credit-to-GDP Gap	0.9	40	100	497	404	23	0.187	0.552	0.023	0.256	0.678	0.813	0.168	0.047
Dom. Credit Growth (yoy)	0.9	58	85	399	502	38	0.309	0.443	0.022	0.240	0.641	0.691	0.176	0.056
Dom. Credit-to-GDP Ratio	0.9	69	51	169	732	72	0.585	0.188	0.019	0.211	0.452	0.415	0.232	0.112
Dom. Credit Gap	0.9	37	104	577	324	19	0.154	0.640	0.018	0.201	0.757	0.846	0.153	0.033
Dom. Credit Growth (4q MA)	0.9	48	93	500	401	30	0.244	0.555	0.017	0.194	0.734	0.756	0.157	0.037
Dom. Credit Growth (6q MA)	0.9	61	72	364	537	51	0.415	0.404	0.015	0.170	0.690	0.585	0.165	0.045
Dom. Credit Growth (qoq)	0.9	46	92	530	371	31	0.252	0.588	0.014	0.153	0.786	0.748	0.148	0.028
Dom. Credit Growth – GDP Growth	0.9	54	70	409	492	53	0.431	0.454	0.009	0.103	0.798	0.569	0.146	0.026
Dom. Credit Growth (8q MA)	0.9	66	57	314	587	66	0.537	0.349	0.009	0.100	0.752	0.463	0.154	0.034

(continued)

Table 3. (Continued)

	μ	Thresh-old	TP	FP	TN	FN	T_1	T_2	Absolute Usefulness	Relative Usefulness	aNtS Ratio	% Predicted	Cond. Prob.	Diff. Prob.
B. Global Variables														
Glo. Credit Gap	0.9	45	113	427	474	10	0.081	0.474	0.040	0.443	0.516	0.919	0.209	0.089
Glo. Credit Growth (qoq)	0.9	60	100	357	544	23	0.187	0.396	0.037	0.412	0.487	0.813	0.219	0.099
Glo. Credit Growth (yoy)	0.9	57	101	365	536	22	0.179	0.405	0.037	0.411	0.493	0.821	0.217	0.097
Glo. Credit Growth (4q MA)	0.9	49	109	448	453	14	0.114	0.497	0.035	0.386	0.561	0.886	0.196	0.076
Glo. Credit Growth (6q MA)	0.9	46	110	467	434	13	0.106	0.518	0.033	0.373	0.580	0.894	0.191	0.071
Glo. Credit Growth (8q MA)	0.9	41	109	509	392	14	0.114	0.565	0.029	0.318	0.637	0.886	0.176	0.056
Glo. Credit-to-GDP Ratio	0.9	75	44	100	801	79	0.642	0.111	0.021	0.229	0.310	0.358	0.306	0.185
Glo. Credit-to-GDP Gap	0.9	37	105	571	330	18	0.146	0.634	0.019	0.216	0.742	0.854	0.155	0.035
Glo. Credit Growth – Glo. GDP Growth	0.9	83	46	161	740	77	0.626	0.179	0.016	0.178	0.478	0.374	0.222	0.102

Notes: The table shows results for the evaluation of individual indicator variables using the signaling approach (see section 3.2 for a detailed description). The preference parameter of $\mu = 0.9$ indicates that policymakers have a strong preference for the detection of crises (i.e., avoiding type I errors) over the avoidance of false alarms (i.e., type II errors). The optimal threshold is calculated as the one that maximizes the relative usefulness and gives the percentile of the country-specific distribution at which the respective indicator issues a warning. The columns of the table report the number of observations where the indicator issues a warning that is indeed followed by a banking crisis seven to twelve quarters ahead (TP); where the indicator issues a warning that is not followed by a crisis (FP); where the indicator issues no warning and there is no crisis seven to twelve quarters ahead (TN); and where the indicator issues no warning although there is a crisis coming (FN). Furthermore, the table reports the fraction of type I errors $T_1 = FN/(TP + FN)$, the fraction of type II errors $T_2 = FP/(FP + TN)$, the absolute and the relative usefulness (see section 3.2 for details), the adjusted noise-to-signal ratio (i.e., the ratio of false signals measured as a proportion of months where false signals could have been issued to good signals as a proportion of months where good signals could have been issued, or $(FP/(FP + TN))/(TP/(TP + FN))$, the percentage of crises correctly predicted (% Predicted), the probability of a crisis conditional on a signal being issued (Cond. Prob.), and the difference between the conditional and the unconditional probability of a crisis (Diff. Prob.). The domestic and the global variables are ranked in terms of relative usefulness.

on a signal being issued (Cond. Prob.), and the difference between the conditional and the unconditional probability of a vulnerable state (Diff. Prob.).

Among the domestic indicators, indeed, the credit-to-GDP gap performs best in the sense that it generates the highest relative usefulness. This is consistent with findings by Drehmann, Borio, and Tsatsaronis (2011) for a different set of countries and in line with the approach taken in the Basel III/ CRD IV framework. The credit-to-GDP gap issues a warning signal whenever it is above the 40th percentile of its country-specific distribution and achieves 25.6 percent of the usefulness a policymaker would gain from a perfectly performing model. Other transformations of the credit variables that perform relatively well are annual credit growth, the credit-to-GDP ratio, and the credit gap (defined as the deviation of the stock of credit from its long-term trend).

Interestingly, global credit variables seem to outperform domestic credit variables in terms of usefulness for predicting vulnerabilities in the domestic banking sector. Panel B of table 3 shows that these indicators usually exert a higher relative usefulness, exert a lower adjusted noise-to-signal ratio, and are able to predict a larger share of the vulnerable states in our sample. In an increasingly integrated economy, vulnerabilities that develop at a global level potentially transmit to countries around the world. Hence, focusing on the development of domestic credit variables might not be sufficient, and the calibration of CCyB rates should also account for global developments. This reasoning is, to some extent, already reflected in the Basel III/CRD IV framework, as the institution-specific CCyB rate is calculated as a weighted average of CCyB rates in countries to which the bank has exposures.

The evaluation of the predictive abilities of global variables is subject to a caveat: As these variables do not vary across countries, and as most countries had a crisis starting in 2008, the good performance of these variables can in part be explained by a clustering of crisis episodes within the same year. That is, indicators based on global credit variables correctly predicted the current crisis in several of our sample countries, which puts the higher usefulness of global as compared with domestic variables in a perspective. However, the current crisis is certainly one of the best examples of a non-domestic vulnerability that spread to banking systems around the world. Thus, if the aim of the CCyB is to increase the resilience of the banking system, it appears to be beneficial to take into account both domestic and global developments.

4.1.2 Multivariate Models Including Other Macrofinancial Indicators

While the signaling approach is a simple and useful way to assess the predictive abilities of individual indicators, a multivariate framework has the advantage of being able to assess the joint performance of several indicators. We therefore estimate simple logit models including several of the individual credit variables as well as other macrofinancial indicators and assess their performance and usefulness.

Results for these models are presented in table 4. Again, we start by considering only the domestic variables and focus on credit growth and the credit-to-GDP gap, as these variables performed well in section 4.1.1 and play a prominent role in the Basel III/CRD IV framework. Credit growth seems to dominate the credit-to-GDP gap, which is statistically not significant, in this simple model. Next, we gradually include the global credit variables, interactions between growth and leverage on the domestic and the global level as well as interactions between the domestic and the global variables.[8] The predictive power of the model improves with each step.[9]

[8]We orthogonalize interaction terms with first-order predictors in order to avoid problems of multicollinearity (see, e.g., Little, Bovaird, and Widaman 2006). In particular, when interacting two variables X and Y, we first form the simple product $X \times Y$ and then regress it on the original variables: $X \times Y = \alpha + \beta_1 \times X + \beta_2 \times Y + \epsilon$. We then take the residual from this regression—ϵ, which is orthogonal to X and Y—to represent the interaction between the two original variables. Variance inflation factors (VIFs) smaller than ten for all variables indicate that we are able to get rid of multicollinearity problems in this way.

[9]Note that the interpretation of interaction effects in logit models is cumbersome. As pointed out by Ai and Norton (2003), the interaction effect is conditional on the independent variables (unlike interaction effects in linear models) and may have different signs for different values of the covariates. Moreover, the statistical significance of these effects cannot be evaluated with a simple t-test, but should be evaluated for each observation separately. Doing so allows us to conclude that for most observations only the Interaction(GC1 $\times$ GC2) is significantly positive, while the other interactions are insignificant (although, e.g., the Interaction(DC2 $\times$ GC2) has a significantly negative sign in the regression itself).

Table 4. Multivariate Models

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5	(6) Model 6	(7) Model 7
Dom. Credit Growth (DC1)	6.38**	2.66	1.61	3.93	1.54	−0.16	0.70
	(2.59)	(2.95)	(2.75)	(3.27)	(3.73)	(4.78)	(3.60)
Dom. Credit-to-GDP Gap (DC2)	3.01	2.71	3.70	8.55***	12.98***	13.24***	20.04***
	(1.93)	(2.80)	(2.66)	(2.27)	(2.27)	(3.19)	(2.68)
Interaction(DC1 × DC2)			26.42	55.68**	55.12**	83.05**	53.94
			(21.83)	(22.77)	(22.35)	(38.97)	(35.03)
Glo. Credit Growth (GCI)		16.71***	16.07***	29.01***	25.99***	19.72*	6.72
		(4.26)	(4.80)	(5.61)	(8.88)	(11.12)	(12.82)
Glo. Credit-to-GDP Gap (GC2)		1.96	−2.74	6.74	12.19	26.84**	41.15***
		(7.67)	(6.57)	(6.67)	(8.89)	(11.72)	(15.31)
Interaction(GC1 × GC2)			391.54**	−486.53*	−324.72	−472.64	−973.05***
			(188.05)	(258.07)	(305.59)	(312.51)	(317.56)
Interaction(DC1 × GC1)			45.98	−56.67	−28.48	−129.64	34.59
			(75.98)	(56.65)	(68.99)	(124.41)	(128.36)
Interaction(DC2 × GC2)			−239.65***	−417.35***	−472.20***	−410.73***	−582.20***
			(49.73)	(67.99)	(91.07)	(100.92)	(109.21)
GDP Growth					19.64	41.84	30.80
					(18.97)	(26.08)	(27.05)
Inflation					−29.04**	−10.04	14.19
					(11.73)	(12.23)	(12.18)
Equity Price Growth					−1.01	−0.38	−0.15
					(1.10)	(1.14)	(1.35)
House Price Growth					16.73***	19.80***	18.05***
					(5.40)	(5.56)	(5.35)

(*continued*)

Table 4. (Continued)

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5	(6) Model 6	(7) Model 7
Global GDP Growth					−10.24	−10.58	−9.88
					(12.62)	(13.68)	(13.79)
Global Equity Price Growth					7.39	7.61	7.78
					(4.80)	(5.42)	(6.12)
Global House Price Growth					16.29	14.97	30.09
					(18.34)	(20.67)	(22.55)
Banking-Sector Capitalization							−136.85***
							(39.63)
Banking-Sector Profitability							324.89***
							(76.45)
Country Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,220	1,220	1,220	919	919	756	756
Pseudo R-Squared	0.0894	0.108	0.133	0.210	0.278	0.272	0.336
AUROC	0.710	0.733	0.780	0.824	0.865	0.846	0.892
Standard Error	0.0266	0.0232	0.0185	0.0195	0.0160	0.0165	0.0157

Notes: The table shows estimation results for multivariate logit models, where the dependent variable is set to 1, seven to twelve quarters preceding a banking crisis in a respective country. Observations for banking crises and six quarters following banking crises are omitted, while other dependent variable observations are set to 0. All regressions include country-specific dummy variables to account for unobserved heterogeneity across countries. Robust standard errors adjusted for clustering at the quarterly level are reported in parentheses. *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively.

In order to compare the models' predictive abilities with those of the individual indicators, we once more apply the signaling approach by translating the predicted probabilities into country-specific percentiles and determining the optimal threshold for the issuance of warnings as the one that maximizes the relative usefulness of the model (see section 3.2). Table 5 shows that the relative usefulness of the domestic model is 0.236, which is lower than that of the best individual indicators. However, the stepwise inclusion of the remaining variables improves the usefulness, so that model 3 surpasses the best domestic as well as the best global indicators in terms of relative usefulness. This indicates the benefits of a multivariate framework as compared with single indicators. We will elaborate more on these benefits by taking into account not only credit variables but also other variables that might affect the stability of the banking sector.

Models 4–7 provide the estimation results for the extended models. The sample size is somewhat smaller than in models 1–3, as the data is not available for all variables across the whole period (see table 1). In order to make results comparable, model 4 reestimates model 3 on the reduced sample. The most striking difference between the two regressions is the coefficient for the domestic credit-to-GDP gap, which turns significant in the reduced sample. This indicates that any evaluation depends on the respective sample and should make policymakers cautious when generalizing findings from a particular sample of countries. However, the predictive abilities of the models are quite impressive. For example, model 5, which we refer to as our *benchmark model*, achieves 60.3 percent of the usefulness of a perfectly performing model and thus outperforms any individual indicator. The area under the ROC curve for this model is equal to 0.865, indicating a good predictive ability of the model for a wide range of policymakers' preference parameters (see figure 3 for an illustration of the ROC curve for our benchmark model).[10]

Overall, we find that the credit variables are indeed among the most important predictors of vulnerable states of the economy. However, both model fit and model performance increase significantly

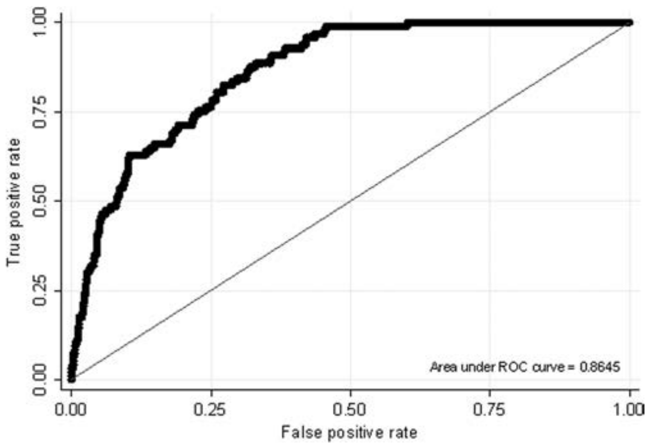
[10]In contrast to the individual indicators and most of the credit models, the extended models perform well also for lower values of the preference parameter μ, which we see as another advantage of these models. Results for other values of μ are available upon request.

Table 5. Model Evaluation

	μ	Thresh-old	TP	FP	TN	FN	T_1	T_2	Absolute Usefulness	Relative Usefulness	aNtS Ratio	% Predicted	Cond. Prob.	Diff. Prob.
Model 1	0.9	48	97	489	427	28	0.224	0.534	0.021	0.236	0.688	0.776	0.166	0.045
Model 2	0.9	43	114	525	391	11	0.088	0.573	0.030	0.336	0.628	0.912	0.178	0.058
Model 3	0.9	56	108	333	583	17	0.136	0.364	0.045	0.497	0.421	0.864	0.245	0.125
Model 4	0.9	67	71	174	501	26	0.268	0.258	0.041	0.456	0.352	0.732	0.290	0.164
Model 5	0.9	63	92	231	444	5	0.052	0.342	0.054	0.603	0.361	0.948	0.285	0.159
Model 6	0.9	63	66	179	408	6	0.083	0.305	0.051	0.595	0.333	0.917	0.269	0.160
Model 7	0.9	67	64	163	424	8	0.111	0.278	0.051	0.596	0.312	0.889	0.282	0.173
Model R1	0.9	42	91	417	366	6	0.062	0.533	0.036	0.401	0.568	0.938	0.179	0.069
Model R2	0.9	65	81	251	532	16	0.165	0.321	0.045	0.503	0.384	0.835	0.244	0.134
Model R3	0.9	69	34	224	451	14	0.292	0.332	0.032	0.357	0.468	0.708	0.132	0.065

Notes: The table shows results for the evaluation of the multivariate models presented in tables 4 and 7. As for the individual indicators, we apply the signaling approach by transforming predicted probabilities into country-specific percentiles. The preference parameter of $\mu = 0.9$ indicates that a policymaker has a strong preference for the detection of crises (i.e., avoiding type I errors) over the avoidance of false alarms (i.e., type II errors). The optimal threshold is calculated as the one that maximizes the relative usefulness and gives the percentile of the country-specific distribution at which the respective indicator issues a warning. The columns of the table report the number of observations where the indicator issues a warning that is indeed followed by a banking crisis seven to twelve quarters ahead (TP); where the indicator issues a warning that is not followed by a crisis (FP); where the indicator issues no warning and there is no crisis seven to twelve quarters ahead (TN); and where the indicator issues no warning although there is a crisis coming (FN). Furthermore, the table reports the fraction of type I errors $T_1 = FN/(TP + FN)$, the fraction of type II errors $T_2 = FP/(FP + TN)$, the absolute and the relative usefulness (see section 3.2 for details), the adjusted noise-to-signal ratio (i.e., the ratio of false signals measured as a proportion of months where false signals could have been issued to good signals as a proportion of months where good signals could have been issued, or $(FP/(FP + TN))/(TP/(TP + FN))$), the percentage of crises correctly predicted (% Predicted), the probability of a crisis conditional on a signal being issued (Cond. Prob.), and the difference between the conditional and the unconditional probability of a crisis (Diff. Prob.).

Figure 3. ROC Curve for Benchmark Model (Model 5)

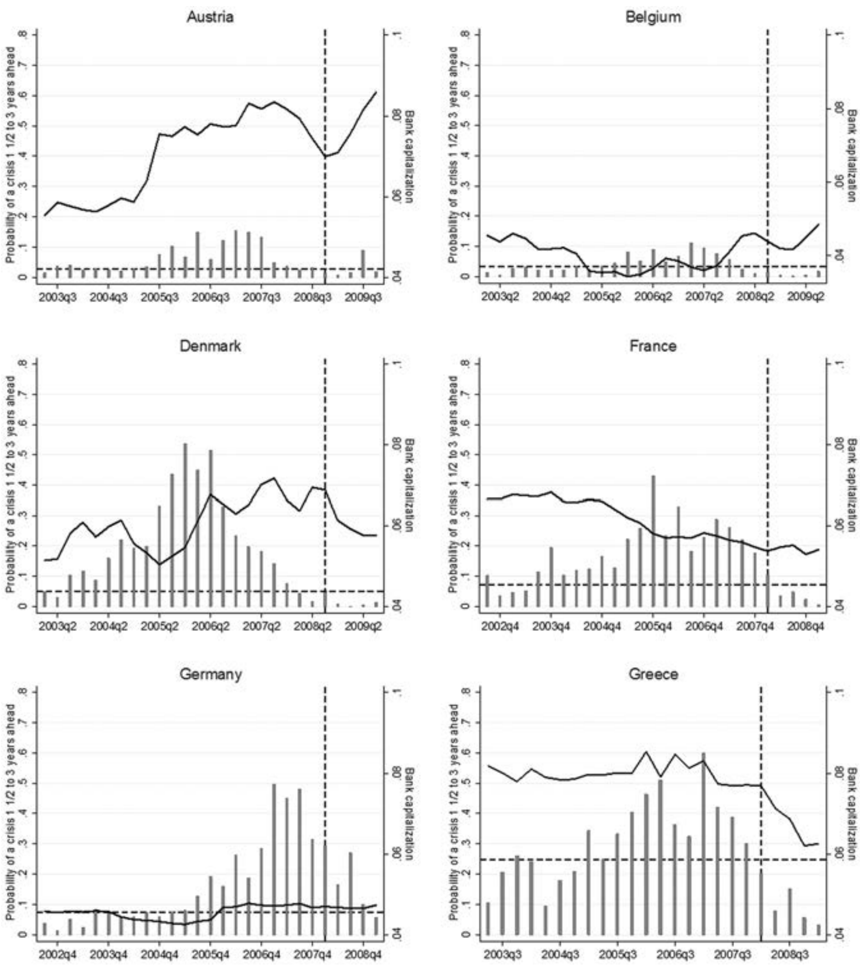


Notes: The figure shows the receiver operating characteristic (ROC) curve for our benchmark model. The area under the ROC curve (AUROC) is equal to 0.8645.

when we include other macrofinancial indicators. For example, the consistently positive coefficient for house price growth indicates that asset price booms promote the buildup of vulnerabilities in the financial sector. This suggests that regulators should keep an eye on these developments instead of focusing exclusively on the development of credit variables. Moreover, model 7 shows that banking-sector variables exert a significant influence on the buildup of financial vulnerabilities. We make the following observations: First, a country is more likely to be in a vulnerable state when aggregate bank capitalization within the country is relatively low. This is a particularly important finding in the context of countercyclical capital buffers, as it indicates that indeed regulators could improve the resilience of the banking system by requiring banks to hold more capital when vulnerabilities build up. Second, we find that future banking crises are more likely when profits in the banking sector are relatively high. As Borio et al. (2010) point out, periods of high bank profitability are typically associated with rapid credit growth, increased risk-taking, and building up of vulnerabilities, which could explain the positive coefficient for the profitability variable preceding banking crises.

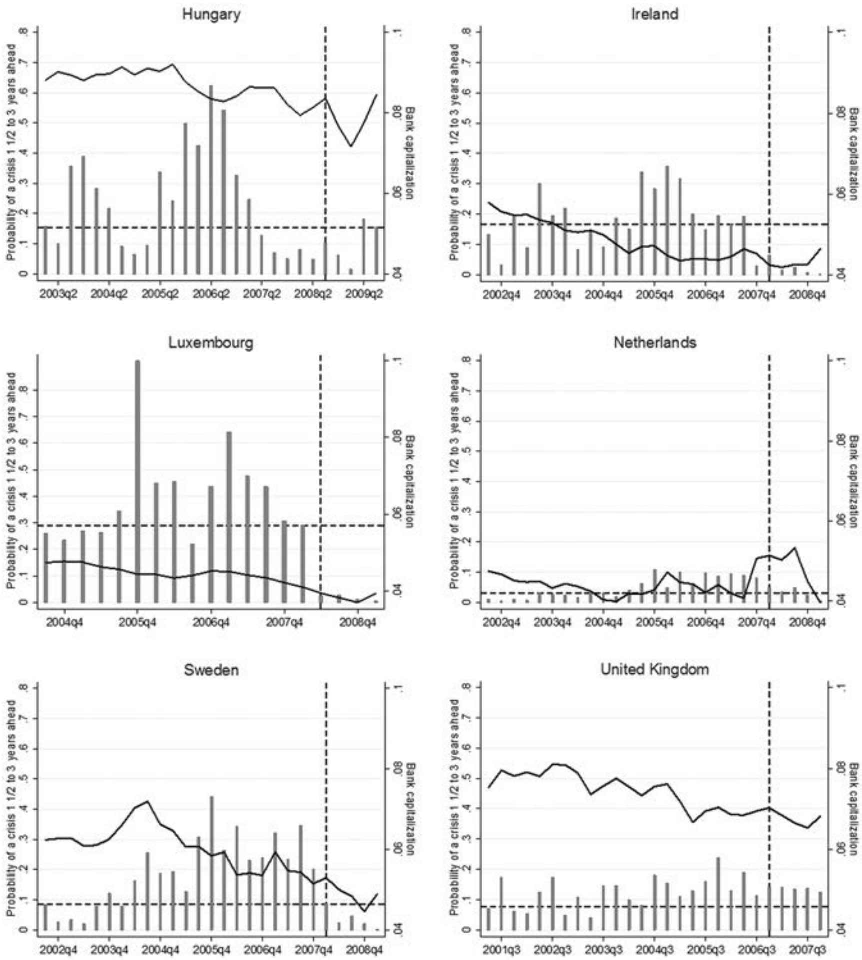
Figure 4 illustrates the relationship between predicted crisis probabilities from our benchmark model (model 5) and actual banking-sector capitalization in the countries that had a banking crisis in 2007/08. Most countries exerted declining or constantly low

Figure 4. Predicted Crisis Probabilities and Banking-Sector Capitalization



(continued)

Figure 4. (Continued)



Notes: The figure plots the predicted probabilities (grey bars) from our benchmark model (model 5 in table 4) around the crises of 2008 in our sample countries (depicted by the dashed vertical lines). The optimal threshold for each country is depicted by the dashed horizontal line. The model issues a warning whenever the predicted probability is above this threshold. The black line shows the development of aggregate capitalization in the banking sector defined as total banking-sector equity over total banking-sector assets.

levels of bank capitalization prior to the crisis, which is consistent with the evidence from model 7. A notable exception is Austria (and to some extent Denmark), where aggregate banking-sector capitalization actually increased prior to the crisis. At the same time, the benchmark model issues a warning already in late-2004/early-2005 in most cases. Hence, if they had relied on this signal, regulators would have had enough time for the activation of the CCyB prior to the crisis—even if we account for an announcement period of twelve months for the CCyB.

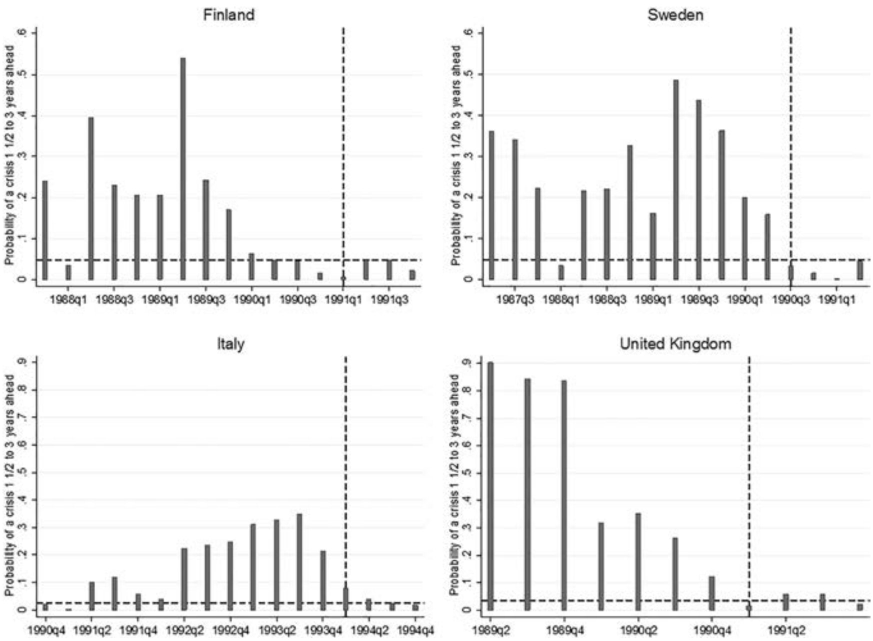
4.2 Out-of-Sample Performance of the Models

Given the objective of the early-warning systems, any assessment should focus on the out-of-sample performance. Moreover, as shown by, e.g., Berg, Borensztein, and Patillo (2005), successful in-sample predictions are much easier to achieve than successful out-of-sample predictions. In order to assess the out-of-sample usefulness of the models, we proceed as follows: First, we consecutively exclude countries that had a banking crisis prior to 2007 from the estimation of the benchmark model. Then, we test whether the model based on the remaining countries is able to predict the crises in the excluded ones.[11]

The results of this exercise are presented in figure 5. The benchmark model signals the banking crises in the Nordic countries well before their onset in the early 1990s. In both Finland and Sweden, the indicator is consistently above the threshold from 1988:Q2 onwards, which is eleven quarters ahead of the crisis for Finland and nine quarters ahead for Sweden. In both cases, banks would have had enough time to build up capital before the crisis if the countercyclical capital buffer had been activated. Similarly, the model issues a warning signal for Italy from 1991:Q2 onwards, eleven quarters ahead of the crisis in 1994. In the United Kingdom, the crisis is relatively close

[11]In principle we could have tried to fit a model to the observations prior to 2007 in order to see whether this model would be able to predict the current crisis. However, as most of the crisis episodes in our sample occur after 2007, and as we particularly want to learn something from these episodes, we prefer the approach described above, i.e., we use the information from the current crisis and check whether it would have been useful for the prediction of past crises.

Figure 5. Out-of-Sample Performance of the Model



Notes: The figure shows results for an out-of-sample evaluation of our benchmark model (model 5 in table 4). We exclude the respective country from the estimation and depict the predicted probabilities (grey bars) from a model based on the remaining countries around the crisis in the excluded country (dashed vertical line). The model issues a warning whenever the predicted probability is higher than the optimal threshold within the country (dashed horizontal line).

to the beginning of the sample period. Yet, in those quarters preceding the crisis of 1991, the benchmark model consistently issues a warning signal. Overall, the benchmark model exhibits strong out-of-sample properties. Information from the current crisis seems to be useful for the prediction of other systemic banking crises in the European Union.

4.3 Robustness Checks

In this section we modify the benchmark model (model 5 of table 4) in several ways in order to further assess the robustness of our results. The results from the robustness analysis are presented in tables 6 and 7.

Table 6. Robustness Checks

	(1) Benchmark	(2) RR	(3) LV	(4) EMU	(5) EU-15	(6) Euro	(7) Interest Rates	(8) Percentiles	(9) Percentiles
Dom. Credit Growth (DC1)	1.54	−10.01**	−0.58	−1.08	1.01	−4.23	2.83	−0.009	−0.000
	(3.73)	(4.71)	(3.28)	(4.29)	(4.57)	(6.77)	(5.86)	(0.008)	(0.010)
Dom. Credit-to-GDP Gap (DC2)	12.98***	23.91***	4.09	15.68***	12.46***	14.70***	37.86***	0.039***	0.050***
	(2.27)	(4.32)	(2.60)	(2.51)	(2.64)	(4.73)	(4.67)	(0.007)	(0.009)
Interaction(DC1 × DC2)	55.12**	35.24	26.75	57.40**	55.10*	138.12*	83.42*	0.018***	0.026***
	(22.35)	(36.30)	(26.96)	(26.67)	(31.69)	(71.79)	(46.34)	(0.005)	(0.005)
Glo. Credit Growth (GCI)	25.99***	32.12***	39.40***	49.67***	18.57**	13.00	108.66***	0.024**	0.018
	(8.88)	(10.77)	(15.26)	(10.14)	(9.42)	(10.23)	(16.80)	(0.010)	(0.011)
Glo. Credit-to-GDP Gap (GC2)	12.19	−8.68	42.49***	−4.67	20.22*	18.52*	−37.62**	−0.033***	0.004
	(8.89)	(10.83)	(15.62)	(12.87)	(11.24)	(10.52)	(17.82)	(0.008)	(0.013)
Interaction(GC1 × GC2)	−324.72	−255.97	1,718.19***	−806.86**	−390.66	−121.35	−1,941.20***	0.011	−0.005
	(305.59)	(320.49)	(587.11)	(345.60)	(296.43)	(301.21)	(527.98)	(0.009)	(0.010)
Interaction(DC1 × GC1)	−28.48	77.07	12.25	−17.42	−240.77**	−134.60	208.01	−0.001	−0.004
	(68.99)	(88.37)	(104.67)	(89.52)	(102.03)	(101.30)	(139.13)	(0.005)	(0.006)
Interaction(DC2 × GC2)	−472.20***	−754.19***	−339.87***	−617.47***	−436.37***	−276.74**	−1,466.07***	−0.042***	−0.062***
	(91.07)	(102.92)	(66.07)	(133.84)	(96.37)	(109.19)	(182.09)	(0.006)	(0.010)
GDP Growth	19.64	2.67	11.35	12.95	18.29	14.00	35.04	0.003	0.001
	(18.97)	(20.61)	(25.41)	(20.57)	(20.65)	(21.40)	(24.89)	(0.006)	(0.006)
Inflation	−29.04**	−3.98	−25.31*	−29.70**	−3.65	15.29	68.72***	−0.001	−0.004
	(11.73)	(11.60)	(15.21)	(12.35)	(10.79)	(10.41)	(20.10)	(0.005)	(0.006)
Equity Price Growth	−1.01	−0.20	−1.46	−1.57	−0.43	−0.14	−1.77	−0.008	−0.009
	(1.10)	(1.49)	(2.11)	(1.16)	(1.26)	(1.51)	(1.91)	(0.006)	(0.006)
House Price Growth	16.73***	14.03***	6.73	21.02***	23.45***	20.71***	20.73**	0.019***	0.015**
	(5.40)	(4.81)	(8.52)	(6.20)	(5.33)	(5.50)	(9.37)	(0.005)	(0.006)

(*continued*)

Table 6. (Continued)

	(1) Benchmark	(2) RR	(3) LV	(4) EMU	(5) EU-15	(6) Euro	(7) Interest Rates	(8) Percentiles	(9) Percentiles
Global GDP Growth	−10.24	−4.87	−40.28*	−8.78	−7.63	−10.67	0.08	−0.004	−0.004
	(12.62)	(12.25)	(20.80)	(12.60)	(12.95)	(15.21)	(11.66)	(0.009)	(0.009)
Global Equity Price	7.39	7.11	21.31***	6.86	8.58*	10.72*	6.47	0.016	0.018*
Growth	(4.80)	(4.56)	(7.62)	(4.99)	(5.07)	(6.14)	(4.61)	(0.010)	(0.010)
Global House Price	16.29	8.29	51.73**	11.56	20.92	27.87	27.49	0.046***	0.054***
Growth	(18.34)	(17.54)	(23.01)	(20.00)	(18.53)	(21.34)	(19.84)	(0.014)	(0.013)
D(EMU)				2.83***					
				(0.62)					
Gov. Bond Yield							−28.69		
							(45.08)		
Money Market Rate							−125.08***		
							(38.55)		
Country Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Observations	919	835	893	919	869	664	862	919	919
Pseudo R-Squared	0.278	0.286	0.442	0.314	0.269	0.270	0.477	0.324	0.372
AUROC	0.865	0.827	0.807	0.887	0.847	0.836	0.942	0.892	0.909
Standard Error	0.0160	0.0207	0.0223	0.0136	0.0172	0.0184	0.0091	0.0126	0.0117

Notes: The table shows robustness checks for our benchmark model (model 5 in table 4). In columns 2 and 3, the dependent variable is based on the crisis definition by Reinhart and Rogoff (2013) and Laeven and Valencia (2008), respectively. Column 4 includes a dummy that is equal to 1 for each quarter in which the respective country is a member of the European Monetary Union (EMU) and column 5 restricts the sample to include only EU-15 countries for which data availability is better than for the rest of our sample countries. Column 6 restricts the sample to include only the countries that are a part of the EMU. In column 7 we include two interest rate variables, the ten-year government bond yield and the three-month interbank lending rate. In columns 8 and 9 we transform all variables into country-specific percentiles before using them in the regression. Robust standard errors adjusted for clustering at the quarterly level are reported in parentheses. *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively.

Table 7. Robustness—Forecast Horizon

	(1) Benchmark 7–12 Quarters	(2) Model R1 1–6 Quarters	(3) Model R2 1–12 Quarters	(4) Model R3 13–20 Quarters
Dom. Credit Growth (DC1)	1.54	−11.63	−3.06	−11.77***
	(3.73)	(8.06)	(4.28)	(4.07)
Dom. Credit-to-GDP Gap (DC2)	12.98***	61.69***	33.92***	16.03***
	(2.27)	(20.72)	(3.91)	(4.08)
Interaction(DC1 × DC2)	55.12**	−18.61	139.78***	−1.36
	(22.35)	(82.99)	(36.95)	(32.76)
Glo. Credit Growth (GCI)	25.99***	113.85***	93.09***	−50.73***
	(8.88)	(18.75)	(12.10)	(14.10)
Glo. Credit-to-GDP Gap (GC2)	12.19	17.90	2.08	28.57**
	(8.89)	(21.61)	(10.95)	(13.06)
Interaction(GC1 × GC2)	−324.72	6,895.95***	2,763.43***	−52.28
	(305.59)	(1,067.10)	(502.18)	(449.06)
Interaction(DC1 × GC1)	−28.48	453.38	223.97**	−606.47***
	(68.99)	(322.98)	(99.32)	(111.78)
Interaction(DC2 × GC2)	−472.20***	−1,947.93***	−1,193.54***	−458.57***
	(91.07)	(565.00)	(151.01)	(79.88)
GDP Growth	19.64	−28.92	−14.58	35.03
	(18.97)	(48.51)	(20.64)	(40.03)
Inflation	−29.04**	77.15***	22.54*	24.42**
	(11.73)	(24.02)	(13.25)	(10.59)

(*continued*)

Table 7. (Continued)

	(1) Benchmark 7–12 Quarters	(2) Model R1 1–6 Quarters	(3) Model R2 1–12 Quarters	(4) Model R3 13–20 Quarters
Equity Price Growth	−1.01	2.14	−0.45	−1.14
	(1.10)	(2.34)	(1.40)	(1.98)
House Price Growth	16.73***	−22.60**	7.29	8.62
	(5.40)	(9.98)	(6.19)	(5.91)
Global GDP Growth	−10.24	−0.39	−2.07	12.86
	(12.62)	(12.63)	(13.10)	(9.72)
Global Equity Price Growth	7.39	14.15***	15.06***	12.53***
	(4.80)	(5.08)	(4.91)	(4.76)
Global House Price Growth	16.29	−60.67**	−32.74	116.13***
	(18.34)	(30.62)	(23.07)	(21.94)
Observations	919	919	919	919
Pseudo R-Squared	0.278	0.781	0.617	0.340
AUROC	0.865	0.726	0.960	0.895
Standard Error	0.0160	0.0179	0.0063	0.0134

Notes: The table shows the results of robustness anlaysis with respect to the forecast horizon. Column 1 reestimates model 5 from table 4, which is referred to as the benchmark model. The dependent variable in this regression is set to 1 seven to twelve quarters preceding a banking crisis in a respective country. In column 2, we replace the dependent variable with a dummy that is equal to 1 one to six quarters before a banking crisis. Similarly, the dependent variable in column 3 equals 1 one to twelve quarters before a banking crisis in the respective country. Finally, the dependent variable in column 4 is equal to 1 thirteen to twenty quarters before the onset of a banking crisis in a respective country. Robust standard errors adjusted for clustering at the quarterly level are reported in parentheses. *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent level, respectively.

First, we check whether our results depend on the definition of the dependent variable. Apart from the ESCB Heads of Research database used in our analysis, the most common definitions of systemic banking crises are provided by Reinhart and Rogoff (2011) and Laeven and Valencia (2012). Although the various databases are broadly consistent with each other, there are some deviations in the timing of crises, as the definition of a systemic event in the banking sector requires a considerable amount of judgment. Columns 2 and 3 show that overall results are relatively similar for all three crisis definitions. Moreover, the area under the ROC curve is also greater than 0.8 for the other two models with the alternative crisis definitions, indicating good predictive abilities of the models.

Second, we include a dummy variable that is equal to 1 for each quarter in which the respective country is a member of the European Monetary Union (EMU). As expected, the coefficient for this dummy variable is positive and significant, as most crises in our sample occur after the establishment of the EMU in 1999 (see column 4). However, the coefficients of the other variables remain largely unaffected by the inclusion of this dummy variable. Furthermore, the results are robust if we restrict the sample to include only countries from the EU-15 (column 5) or only countries that are part of the EMU (column 6).

Third, we augment the model with a money market rate (column 7). The estimated negative coefficient is potentially related to the "great moderation," i.e., the general decline of inflation and money market rates over the sample period. The high R-squared and AUROC indicate that the fit of the model is superior to that of the other models. Despite this, we do not select this model as our benchmark model, as its out-of-sample forecast abilities are inferior to the benchmark model, potentially due to an overfitting problem.

Fourth, following Lo Duca and Peltonen (2013), we transform all variables into country-specific percentiles before using them in the regression. This method can be seen as an alternative way to account for heterogeneity across countries, as differences in levels of indicators between countries vanish for the transformed variables. Columns 8 and 9 show that most of the estimated coefficients have the same sign as in the benchmark model if we use this alternative method.

Finally, we analyze model performance across different forecast horizons (see also Schudel 2013). Specifically, we check how the performance of the benchmark model and the indicator properties of variables change if the time window of the vulnerable state preceding a systemic banking crisis is altered from the seven to twelve quarters used in the standard specifications. Results in table 7 show that although the benchmark model is broadly robust to an alteration of the forecast horizon, the relative importance and the estimated signs of the coefficients tend to vary a bit. Particularly important are the reversed signs for domestic and global credit growth in the model with the thirteen- to twenty-quarters-ahead definition of a vulnerable state and the strong influence of global asset prices in this model (model R3). As shown in table 5, the benchmark model with a forecast horizon of seven to twelve quarters (model 5) provides the highest absolute and relative usefulness measures, followed by the model with a forecast horizon of one to twelve quarters (model R2). The performances of the models with the early (one to six quarters) and late (thirteen to twenty quarters) pre-crisis time horizons in terms of absolute and relative usefulness are broadly similar to, but markedly lower than, that of the benchmark model.

5. Conclusion

As a response to recent financial crises, the Basel III/CRD IV regulatory framework includes a countercyclical capital buffer (CCyB) to increase the resilience of the banking sector and its ability to absorb shocks arising from financial and economic stress. In this context, this paper seeks to provide an early-warning model, which can be used to guide the buildup and release of capital in the banking sector. Given the prominence of private credit variables in the Basel III/CRD IV framework, the paper first examines the evolution of credit variables preceding banking crises in the EU member states and assesses their usefulness in guiding the setting of the CCyB. Furthermore, the paper examines the potential benefits of complementing private credit variables with other macrofinancial and banking-sector indicators in a multivariate logit framework. The evaluation of the policy usefulness of the credit indicators and models follows the methodology applied in Alessi and Detken (2011), Lo Duca and Peltonen (2013), and Sarlin (2013).

The paper finds that, in addition to credit variables, other domestic and global financial factors such as equity and house prices and banking-sector variables help to predict macrofinancial vulnerabilities in EU member states. The importance of global variables in our models concurs with the view that excessive global liquidity is a key driver of financial vulnerabilities and highlights the importance of consistent cross-border and reciprocity arrangements, as in the European macroprudential framework. Furthermore, higher banking-sector capitalization decreases the probability of entering a vulnerable state, which provides a rationale for the implementation of countercyclical measures like the CCyB. Moreover, future banking crises tend to be more likely when profits in the banking sector are relatively high, concurrent with the view that high bank profitability could be associated with rapid credit growth, increased risk-taking, and building up of vulnerabilities. Overall, our findings suggest that policymakers should take a broad approach in their analytical models supporting CCyB policy measures.

Appendix. Data Sources and Technical Estimation Matters

Data Sources

The individual series used in our paper stem from the following original sources: Data on total credit to the private non-financial sector are obtained from the BIS and—for those countries where BIS data is not available—from Eurostat. Information on nominal GDP growth and inflation rates comes from the IMF's International Financial Statistics (IFS). Data on stock prices are obtained from the OECD, while data on house prices are provided by the BIS. Banking-sector variables are obtained from two sources: The OECD provides relatively long series on banking-sector capitalization and profitability on an annual basis that we use in the empirical analysis. Additionally, for illustrative purposes, we use a shorter series of banking-sector capitalization in figure 4 that is available on a quarterly basis and that is obtained from the ECB's Balance Sheet Items (BSI) statistics. Finally, quarterly data on the ten-year government bond yield and the three-month interbank lending rate (money market rate) are obtained from the OECD.

Calculation of the Credit-to-GDP Gap

Following the Basel Committee on Banking Supervision (2010), we use a backward-looking Hodrick-Prescott filter with a smoothing parameter λ of 400,000 to calculate the credit-to-GDP gap. Recommendations in the BCBS Consultative Document are based on a paper by Borio et al. (2010), who find that trends calculated with a λ of 400,000 perform well in picking up the long-term development of private credit. In particular, a λ of 400,000 is consistent with the assumption of credit cycles being four times longer than business cycles if one follows a rule developed by Ravn and Uhlig (2002), which states that the optimal λ of 1,600 for quarterly data should be adjusted by the fourth power of the observation frequency ratio (i.e., if credit cycles are four times longer than business cycles, λ should be equal to $44 \times 1,600 \approx$ 400,000).

Receiver Operating Characteristics (ROC) Curves

In addition to assessing the relative and absolute usefulness of a model, we also employ receiver operating characteristics (ROC) curves and the area under the ROC curve (AUROC) for comparing performance of the early-warning models. The ROC curve shows the tradeoff between the benefits and costs of a certain threshold τ. When two models are compared, the better model has a higher benefit (TP rate (TPR) on the vertical axis) at the same cost (FP rate (FPR) on the horizontal axis).[12] Thus, as each FP rate is associated with a threshold, the measure shows performance over all thresholds or, equivalently, over all preference parameters of the policymaker. The AUROC is computed using trapezoidal approximations and measures the probability that a randomly chosen vulnerable state receives a higher predicted probability than a tranquil period. A perfect ranking has an AUROC equal to 1, whereas a coin toss has an expected AUROC of 0.5.

[12]The TPR (also called sensitivity) gives the ratio of periods where the model correctly issues a warning to all periods where a warning should have been issued, formally $TPR = TP = (TP + FN)$. The FPR (also called specificity) gives the ratio of periods where the model wrongly issues a signal to all periods where no signal should have been issued, formally $FPR = FP = (FP + TN)$. An ideal model would achieve a TPR of 1 (no missed crises) and an FPR of 0 (no false alarms).

References

Ai, C., and E. C. Norton. 2003. "Interaction Terms in Logit and Probit Models." *Economics Letters* 80 (1): 123–29.

Alessi, L., A. Antunes, J. Babecky, S. Baltussen, M. Behn, D. Bonfim, O. Bush et al. 2015. "Comparing Different Early Warning Systems: Results from a Horse Race Competition among Members of the Macroprudential Research Network." Macro-Prudential Research Network, European Central Bank.

Alessi, L., and C. Detken. 2011. "Quasi Real Time Early Warning Indicators for Costly Asset Price Boom/Bust Cycles: A Role for Global Liquidity." *European Journal of Political Economy* 27 (3): 520–33.

———. 2014. "Identifying Excessive Credit Growth and Leverage." ECB Working Paper No. 1723.

Allen, F., and D. Gale. 2009. *Understanding Financial Crises.* Oxford University Press.

Babecky, J., T. Havranek, J. Mateju, M. Rusnák, K. Smidkova, and B. Vasicek. 2014. "Banking, Debt, and Currency Crises in Developed Countries: Stylized Facts and Early Warning Indicators." *Journal of Financial Stability* 15 (December): 1–17.

Barrell, R., E. P. Davis, D. Karim, and I. Liadze. 2010. "Bank Regulation, Property Prices and Early Warning Systems for Banking Crises in OECD countries." *Journal of Banking and Finance* 34 (9): 2255–64.

Basel Committee on Banking Supervision. 2010. "Guidance for National Authorities Operating the Countercyclical Capital Buffer." Bank for International Settlements.

Behn, M., C. Detken, T. Peltonen, and W. Schudel. 2013. "Setting Countercyclical Capital Buffers based on Early Warning Models: Would It Work?" ECB Working Paper No. 1604.

Berg, A., E. Borensztein, and C. Patillo. 2005. "Assessing Early Warning Systems: How Have They Worked in Practice?" *IMF Staff Papers* 52 (3): 462–502.

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

Brunnermeier, M., and M. Oehmke. 2013. "Bubbles, Financial Crises, and Systemic Risk." In *Handbook of the Economics of*

Finance, Vol. 2, Part B, ed. G. M. Constantinides, M. Harris, and R. M. Stulz, 1221–88 (chapter 18). Elsevier.

Bussière, M., and M. Fratzscher. 2006. "Towards a New Early Warning System of Financial Crises." *Journal of International Money and Finance* 25 (6): 953–73.

Caprio, G., and D. Klingebiel. 2003. "Episodes of Systemic and Borderline Financial Crises." World Bank Research Dataset.

Castro, C., A. Estrada, and J. Martínez-Pagés. 2016. "The Countercyclical Capital Buffer in Spain: An Analysis of Key Guiding Indicators." Working Paper No. 1601, Banco de Espana.

Committee on the Global Financial System. 2011. "Global Liquidity — Concept, Measurement and Policy Implications." CGFS Paper No. 45, Bank for International Settlements.

Dembiermont, C., M. Drehmann, and S. Muksakunratana. 2013. "How Much Does the Private Sector Really Borrow? A New Database for Total Credit to the Private Non-financial Sector." *BIS Quarterly Review* (March): 65–81.

Demirgüç-Kunt, A., and E. Detragiache. 1998. "The Determinants of Banking Crises in Developing and Developed Countries." *IMF Staff Papers* 45 (1): 81–109.

———. 2000. "Monitoring Banking Sector Fragility. A Multivariate Logit Approach." *World Bank Economic Review* 14 (2): 287–307.

Detragiache, E., and A. Spilimbergo. 2001. "Crises and Liquidity: Evidence and Interpretation." IMF Working Paper No. 01/2.

Dickey, D. A., and W. A. Fuller. 1979. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association* 74 (366a): 427–31.

Drehmann, M., C. Borio, and K. Tsatsaronis. 2011. "Anchoring Countercyclical Capital Buffers: The Role of Credit Aggregates." *International Journal of Central Banking* 7 (4): 189–240.

———. 2013. "Evaluating Early Warning Indicators of Banking Crises: Satisfying Policy Requirements." BIS Working Paper No. 421.

Drehmann, M., and M. Juselius. 2014. "Evaluating Early Warning Indicators of Banking Crises: Satisfying Policy Requirements." *International Journal of Forecasting* 30 (3): 759–80.

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

Ferrari, S., and M. Pirovano. 2015. “Early Warning Indicators for Banking Crises: A Conditional Moments Approach.”

Frankel, J. A., and A. K. Rose. 1996. “Currency Crashes in Emerging Markets: An Empirical Treatment.” *Journal of International Economics* 41 (3–4): 351–66.

Galati, G., I. Hindrayanto, S. J. Koopman, and M. Vlekke. 2016. “Measuring Financial Cycles in a Model-based Analysis: Empirical Evidence for the United States and the Euro Area.” *Economics Letters* 145 (August): 83–87.

Hanschel, E., and P. Monnin. 2005. “Measuring and Forecasting Stress in the Banking Sector: Evidence from Switzerland.” *BIS Papers* 22 (April): 431–49.

Holopainen, M., and P. Sarlin. 2016. “Toward Robust Early-Warning Models: A Horse Race, Ensembles and Model Uncertainty.” ECB Working Paper No. 1900.

Im, K. S., M. H. Pesaran, and Y. Shin. 2003. “Testing for Unit Roots in Heterogeneous Panels.” *Journal of Econometrics* 115 (1): 53–74.

International Monetary Fund. 2013. “Global Liquidity — Credit and Funding Indicators.” IMF Policy Paper (July).

Ito, Y., T. Kitamura, K. Nakamura, and T. Nakazawa. 2014. “New Financial Activity Indexes: Early Warning System for Financial Imbalances in Japan.” Working Paper No. 14-E-7, Bank of Japan.

Jordà, Ò., M. Schularick, and A. Taylor. 2015. “Leveraged Bubbles.” *Journal of Monetary Economics* 76 (Supplement): S1–S20.

Kaminsky, G., S. Lizondo, and C. M. Reinhart. 1998. “Leading Indicators of Currency Crises.” *IMF Staff Papers* 45 (1): 1–48.

Kaminsky, G. L. 2006. “Currency Crises: Are They All the Same?” *Journal of International Money and Finance* 25 (3): 503–27.

Kaminsky, G. L., and C. M. Reinhart. 1999. “The Twin Crises: The Causes of Banking and Balance-of-Payments Problems.” *American Economic Review* 89 (3): 473–500.

Karim, D., I. Liadze, R. Barrell, and P. Davis. 2013. "Off-balance Sheet Exposures and Banking Crises in OECD Countries." *Journal of Financial Stability* 9 (4): 673–81.

Laeven, L., and F. Valencia. 2008. "Systemic Banking Crises: A New Database." IMF Working Paper No. 08/224.

———. 2010. "Resolution of Banking Crises: The Good, the Bad, and the Ugly." IMF Working Papers No. 10/146.

———. 2012. "Systemic Banking Crises Database: An Update." IMF Working Paper No. 12/163.

Little, T. D., J. A. Bovaird, and K. F. Widaman. 2006. "On the Merits of Orthogonalizing Powered and Product Terms: Implications for Modeling Interactions among Latent Variables." *Structural Equation Modeling* 13 (4): 497–519.

Lo Duca, M., and T. Peltonen. 2013. "Assessing Systemic Risks and Predicting Systemic Events." *Journal of Banking and Finance* 37 (7): 2183–95.

López-Salido, D., J. Stein, and E. Zakrajŝek. 2016. "Credit-Market Sentiment and the Business Cycle." NBER Working Paper No. 21879.

Manasse, P., R. Savona, and M. Vezzoli. 2016. "Danger Zones for Banking Crises in Emerging Markets." *International Journal of Finance and Economics* 21 (4): 360–81.

Park, J. Y., and P. C. Phillips. 2000. "Nonstationary Binary Choice." *Econometrica* 68 (5): 1249–80.

Ravn, M. O., and H. Uhlig. 2002. "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations." *Review of Economics and Statistics* 84 (2): 371–76.

Reinhart, C. M., and K. S. Rogoff. 2011. "From Financial Crash to Debt Crisis." *American Economic Review* 101 (5): 1676–1706.

———. 2013. "Banking Crises: An Equal Opportunity Menace." *Journal of Banking and Finance* 37 (11): 4557–73.

Repullo, R., and J. Saurina. 2011. "The Countercyclical Capital Buffer of Basel III: A Critical Assessment." CEPR Discussion Paper No. DP8304.

Sarlin, P. 2013. "On Policymakers' Loss Functions and the Evaluation of Early Warning Systems." *Economics Letters* 119 (1): 1–7.

Schudel, W. 2013. "Shifting Horizons: Assessing Macro Trends Before, During, and Following Systemic Banking Crises." ECB Working Paper No. 1766.

Schularick, M., and A. M. Taylor. 2012. "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870–2008." *American Economic Review* 102 (2): 1029–69.

Schüler, Y., P. Hiebert, and T. Peltonen. 2015. "Characterising the Financial Cycle: A Multivariate and Time-varying Approach." ECB Working Paper No. 1846.

———. 2017. "Coherent Financial Cycles for G-7 Countries: Why Extending Credit Can Be an Asset." ESRB Working Paper No. 43.

Seidler, J., and A. Gersl. 2011. "Credit Growth and Capital Buffers: Empirical Evidence from Central and Eastern European Countries." Research and Policy Note No. 2/2011, Czech National Bank.

Yeyati, E. L., and U. Panizza. 2011. "The Elusive Costs of Sovereign Defaults." *Journal of Development Economics* 94 (1): 95–105.

The Great Globalization and Changing Inflation Dynamics*

Chengsi Zhang
School of Finance and China Financial Policy Research Center,
Renmin University of China

This paper investigates the changing impact of economic globalization on U.S. inflation dynamics during the period from 1980 to 2012. We develop an extended New Keynesian Phillips curve incorporating richer dynamics and foreign economic slack from microfoundations. Using unknown structural break tests, we identify a significant structural change in the early 1990s in the model, after which domestic inflation responds more strikingly to the foreign than to the domestic output gap. The finding indicates that it is plausible for the Federal Reserve to augment its policy analysis framework with globalization factors in the globalized world.

JEL Codes: E31, E37, E52, E58, C22.

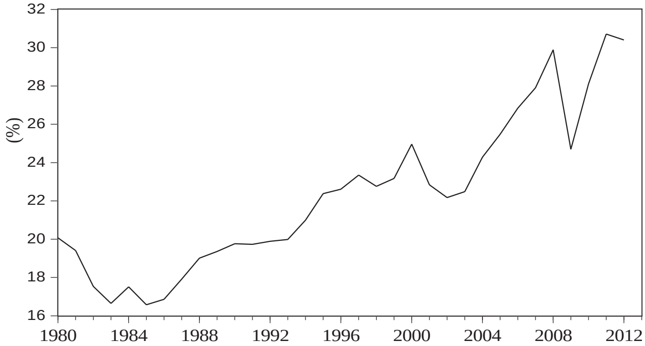
1. Introduction

The United States is quite an open economy today, and its trading partners' economic performance plays an important role in the slowly solidifying recent cyclical recovery. The rising economic globalization also brings an increasingly complex set of policy challenges. One of the leading challenges is the changing nature of inflation dynamics in the process of the current great globalization.[1] The past

*The research was supported by the Ministry of Education of China (Grant No. 16JJD790057). The author is grateful to the anonymous referee for helpful comments on an earlier version of the paper. All errors and omissions are the author's sole responsibility. Author e-mail: zhangcs@ruc.edu.cn.

[1]The current age of globalization is regarded as the second great wave of globalization of international trade (and capital flows). The first wave of globalization occurred from the mid-1850s to the First World War; see O'Rourke and Williamson (1999) and Mishkin (2006) for the insight analyses of the first and the second waves of the great globalization, respectively.

Figure 1. U.S. Economic Globalization (exports plus imports of goods and services, % of GDP): 1980–2012



Source: U.S. Department of Commerce: Bureau of Economic Analysis, and the author's calculations.

decades saw a marked fall in U.S. inflation, also associated with a distinct increase in economic globalization. As figure 1 attests, the level of economic globalization in the United States, measured by total trade as percentage of nominal gross domestic product (GDP), has become increasingly more open over the past three decades, at a pace that has further intensified since the mid-1990s. The rising globalization potentially brings a new challenge for monetary authority: the central bank could cost the nation a dangerous lurch into secular deflation or unexpected high inflation if they fail to grasp profound global changes at play and their implications for domestic inflation developments.

This concern has ignited a growing number of studies investigating the effects of globalization on inflation, but the conclusions remain very different. The subject of the active debate in the literature is whether globalization has reduced the sensitivity of inflation to domestic economic slack while increasing the response of inflation to foreign economic slack. One branch of the literature believes that globalization has indeed led to a decline in the sensitivity of inflation to domestic factors. They argue that the integration of China and other lower-cost producers in world production networks increases competition and induces downward pressure on wages and import prices in industrial countries, which in turn remarkably subdues domestic inflation despite notable rises in commodity prices and a broadly accommodating monetary policy stance across the major currency areas. Gamber and Hung (2001), International Monetary Fund (2006), Pain, Koske, and Sollie (2006), Borio and Filardo (2007), Pehnelt (2007), Auer, Degen, and Fischer (2010), and Guerrieri, Gust, and López-Salido (2010) appear to support this argument.

In contrast, the empirical results for the euro area in Calza (2008), for Organization for Economic Cooperation and Development (OECD) economies in Ihrig et al. (2010), for the United States in Milani (2012), and for France, the United Kingdom, and the United States in Lopez-Villavicencio and Saglio (2014) show that globalization has little impact on the inflation process in major industrial countries. Still, other studies offer alternative explanations for the changing inflation performance in the industrial countries, including improved cyclical conditions (Buiter 2000), systematic improvements in monetary policy (Ball and Moffitt 2001; Kamin, Marazzi, and Schindler 2004; Roberts 2006; Mishkin 2007), or simply good luck (Stock and Watson 2007).

These conflicting findings may reflect the fact that the integration of emerging countries into the global economy can bring interconnecting and two-way impacts on the inflation process. On the one hand, higher demand may drive up prices for energy, raw materials, and general commodities, which will eventually reflect in overall price inflation. On the other hand, an influx of lower-cost labor, products, and services into the world market can also drive prices downwards. This two-way impact may also explain why the globalization–inflation relationship remains a "puzzle" (Temple 2002) when different pools of countries are considered as in Romer (1993) and Gruben and McLeod (2004).

The discrepancy in the literature also reflects the importance of reasonable model specification used in analyzing the globalization–inflation connection. For example, Borio and Filardo (2007) specify a stylized Phillips curve with lagged real domestic and foreign output gaps and find that the globalization factor (captured by foreign output gap) is a key determinant of domestic inflation. Instead of using lagged output gaps, Ihrig et al. (2010) modify Borio and Filardo's model specification by using contemporaneous output gaps (and making changes in the country weights, among other things), and the results appear to change dramatically. Ball (2006) and Mishkin

(2009), citing the empirical results in Ihrig et al. (2010), argue that there is little reason to think that globalization has changed the structure of inflation dynamics. In addition, the simulation results of the theoretical experiments based on a three-equation model of monetary transmission mechanism in Woodford (2009) also show that globalization does not weaken the ability of national central banks to control the dynamics of inflation.

The aforementioned studies provide useful contributions to the understanding of inflation dynamics in a globalized world. Two important issues, however, are underexamined in the literature. The first is the structural specification of an inflation dynamics model with globalization factors. The commonly used framework in the literature of globalization and inflation is a reduced-form model of a traditional backward-looking Phillips curve (e.g., Borio and Filardo 2007; Ihrig et al. 2010), which maintains the virtue of simplicity but neglects important microfoundations of the staggered-price-setting mechanism. According to modern monetary theory, estimated reduced-form inflation equations provide very little information regarding the role of global factors as determinants of inflation (Galí 2010). The backward-looking Phillips-curve framework also brings arbitrariness in detailed model specifications, which leads to difficulties in discriminating different results in the relevant studies.

The structural monetary models of an open economy developed in the seminal works of Clarida, Galí, and Gertler (2002), Woodford (2003), and Galí and Monacelli (2005) provide an important benchmark for structural modeling of inflation dynamics. Milani (2010, 2012) estimates the structural New Keynesian models to test whether globalization has changed the behavior of inflation dynamics in the United States and other G-7 economies. These structural models, however, emphasize the pure forward-looking behavior of inflation, which may not be able to capture the observed inflation inertia in the United States (and many other countries). Even though there are some studies that estimate hybrid New Keynesian Phillips-curve models with both forward- and backward-looking components, the backward-looking behavior is only captured by a stylized one lag. In this paper, we will extend the recent pricing behavior of the backward-looking firms to incorporate a weighted process of past inflation instead of a stylized one lag. We will also show that an extended model with both forward- and backward-looking inflation components and a foreign economic slack can be rationalized with a slight modification in a similar microeconomic framework to the literature.

A second, equally important but underexamined issue is potential structural changes in the impact of globalization on inflation dynamics. Over the long span of the past three decades in the United States, for example, the level of economic globalization has changed (increased) and hence the machinery of its economic interaction with the world economy may have experienced considerable changes. The rising globalization may make domestic inflation more responsive to foreign than to domestic economic slack via different channels. In a globalized world, foreign goods and services provide alternative choices for domestic consumers, so there will be less pressure for domestic prices to rise (Mishkin 2009). Rising globalization may also make the domestic firms more sensitive now than before to foreign economic performance in their price-setting process so that global factors will eventually have an impact on domestic inflation. Through either channel, the impact of globalization on domestic inflation may change over time. Borio and Filardo (2007) have noticed a possible impact of structural breaks on the effect of globalization inflation, but their analysis is largely based on informal break tests. Therefore, formal structural break tests are warranted in the relevant studies.

This paper attempts to fill the above two voids and focuses on the changing impact of globalization on inflation dynamics in the United States, linking the results to broader debates in the academic literature as well as policy implications. To this end, we develop an extended New Keynesian Phillips curve (NKPC) from an open-economy version of an extended Calvo's (1983) sticky price model with microeconomic foundations relating to the dynamic stochastic general equilibrium model. The extended NKPC incorporates both inflation expectations and inflation inertia, and includes foreign economic slack to provide a channel through which globalization may alter the dynamic response of inflation to domestic excess demand. Based on this model, we carry out unknown structural break tests over the period of 1980 to 2012 and find a significant structural break in the U.S. inflation dynamics in the early 1990s, after which an increase in globalization generates a larger increase in the response of inflation to the foreign than the domestic output gap.

The rest of the paper is organized as follows: Section 2 presents the baseline model and describes its connection with the microfoundation of sticky prices (with detailed model development presented in the appendix); section 3 describes the data used in empirical work and some stylized facts of globalization and inflation in the United States; section 4 discusses the econometric issues related to the empirical estimations and provides empirical results of the underlying model, followed by relevant implications explored in section 5; section 6 concludes the paper.

2. The Model

In this section, we construct a dynamic, microfounded inflation dynamics model for an open economy with sticky prices. Our model is a small but presumably important extension of recent developments in the open-economy dynamic stochastic general equilibrium (DSGE) literature pioneered by Obstfeld and Rogoff (1995, 2000), Clarida, Galí, and Gertler (2002), Smets and Wouters (2002), and Galí and Monacelli (2005). The households' consumption and savings decisions are in the spirit of Galí and Monacelli (2005), in which a representative household seeks to maximize a utility function with a composite consumption composed of both domestic goods and imported goods.

For the domestic monopolistically competitive goods market, we lay out an extended framework of Calvo (1983) that can be easily rationalized in terms of sticky price setting of backward-looking firms in the closed-economy models. In particular, for the pricing behavior of the domestic firms, we assume an economic environment similar to Calvo's (1983) model, in which firms are able to revise their prices in any given period with a fixed probability $(1 - \theta)$. In addition, we assume both "forward-" and "backward-looking" firms coexist in the economy with a proportion of ω and $(1 - \omega)$, respectively. Further, we extend the rule of the recent pricing behavior of the backward-looking firms to incorporate a weighted process of past inflation, instead of stylized one lag of inflation inertia.

Specifically, based on the regular assumptions in Calvo's model and log-linear approximations, it is possible to obtain the (log) aggregate price level by

$$p_t = \theta p_{t-1} + (1-\theta)p_t^{new}, \tag{1}$$

where p_t^{new} is the new price set in period t. Let p_t^f be the price set by forward-looking firms and p_t^b be the price set by backward-looking firms at time t. The new price (relative to the aggregate price) can be expressed as a convex combination of p_t^f and p_t^b, viz.

$$p_t^{new} - p_t = \omega(p_t^f - p_t) + (1-\omega)(p_t^b - p_t). \tag{2}$$

Next, following Woodford (2003), the pricing behavior of the forward-looking firms can be written as

$$p_t^f - p_t = \theta\varsigma \sum_{s=0}^{\infty} (\theta\varsigma)^s E_t \pi_{t+s+1} + (1-\theta\varsigma) \sum_{s=0}^{\infty} (\theta\varsigma)^s E_t mc_{t+s}, \tag{3}$$

where ς denotes a subjective discount factor, π_t denotes inflation rate, and mc_t is the real marginal cost of a typical domestic firm producing differentiated goods with a linear technology. Iterating (3) gives

$$p_t^f - p_t = \theta\varsigma E_t \pi_{t+1} + (1-\theta\varsigma)mc_t + \theta\varsigma E_t(p_{t+1}^f - p_{t+1}). \tag{4}$$

Assume a rule of thumb in the pricing setting, viz.

$$p_t^b - p_t = p_{t-1}^{new} - p_t + \pi_{t-1}. \tag{5}$$

As emphasized in the literature, this is an elegant innovation in that the backward-looking firms can now set their prices to the average price determined in the most recent price adjustments with a correction for inflation.

However, inflation inertia in (5) is confined to one single lag, which may neglect the importance of other historical inflation in predicting current inflation. In particular, if we interpret one period as being short, the backward-looking agents are likely to take more than one period to fully respond to changes in actual inflation. Therefore, it would appear reasonable to replace π_{t-1} in (5) with a weighted average of inflation over several periods in the past. Importantly, this replacement can effectively mitigate a serial correlation problem in empirical analysis. As such, we extend (5) in the following process:

$$p_t^b - p_t = (p_{t-1}^{new} - p_t) + \pi_{t-1} + \rho(L)\Delta\pi_{t-1}, \tag{6}$$

where $\rho(L) = \rho_1 + \rho_2 L + \cdots + \rho_m L^{m-1}$ is a polynomial in the lag operator. In practice, m may be specified by utilizing appropriate diagnostic tests (e.g., standard information criteria).

Combining (1)–(6), it can be shown that the dynamics of domestic inflation in terms of real marginal cost are described by an equation analogous to the one associated with a closed economy, viz.

$$\pi_t = \gamma_e E_t \pi_{t+1} + \gamma_b \pi_{t-1} + \sum_{i=1}^{m-1} \alpha_i \Delta \pi_{t-i} + \lambda mc_t, \tag{7}$$

where coefficients of (7) are functions of the deeper parameters in (1)–(6).

As shown in the appendix of this paper, combining equation (7) with the equations depicting the representative household seeking to maximize a utility function with a composite consumption composed of both domestic goods and imported goods, we can obtain an open-economy generalization of the extended New Keynesian Phillips curve (NKPC) that incorporates an extended inflation dynamics and world excess demand, viz.

$$\pi_t = c + \gamma_e E_t \pi_{t+1} + \gamma_b \pi_{t-1} + \sum_{i=1}^{m-1} \alpha_i \Delta \pi_{t-i} + \delta_d \hat{y}_t + \delta_f \hat{y}_t^* + \lambda_t, \tag{8}$$

where $\hat{y}_t$ and $\hat{y}_t^*$ denote domestic and foreign real output gaps, λ_t is an error term, and other coefficients are functions of structural parameters in the DSGE system illustrated in the appendix.

Note that model (8) introduces the globalization factor and more inflation dynamics (i.e., $\sum_{i=1}^{m-1} \alpha_i \Delta \pi_{t-i}$) into an extended NKPC model and provides a channel through which globalization may alter the dynamic response of inflation to domestic demand. If the impact of the world excess demand on inflation is trivial (insignificant), the model then reduces to a closed-economy version of NKPC. The inclusion of additional lag terms in the model is particularly important in order to obtain valid results in empirical estimations, since the stylized specification of the conventional NKPC model generally has a serial correlation problem (Zhang, Osborn, and Kim 2008). The possible presence of serial correlation is crucial for the choice of valid instruments for GMM estimations of the NKPC models, since all lags of the dependent variable are invalid instruments in the presence of autoregressive serial correlation. Since lags of inflation are typically employed as instruments for estimation of the NKPC, the consistency of these estimates depends on the lack of such serial correlation.

3. The Data

The baseline estimation of model (8) involves series for domestic inflation π_t; inflation expectations $E_t\pi_{t+1}$; a measure of the domestic output gap; and a measure of the foreign output gap $\hat{y}_t^*$. As will be discussed in section 4, the estimation of model (8) also involves effective federal funds rate and growth rate of monetary aggregate (i.e., M2) as instrumental variables. The raw (seasonally adjusted) data were obtained from the Federal Reserve Economic Data (FRED) database of the Federal Reserve Bank of St. Louis. The final series used in empirical estimations are stationary (confirmed by conventional unit-root tests). The sample size in our empirical estimations spans the first quarter of 1980 to the last quarter of 2012 (i.e., 1980:Q1–2012:Q4) dictated by availability of trade data. Available monthly data on consumer price index, federal funds rate, and M2 are transformed using end-of-quarter observations as the corresponding quarter values to avoid inducing serial correlation in the final data set.

Several issues of the data deserve further explanations. First, U.S. inflation is measured by the quarterly year-on-year growth rate of the consumer price index (CPI). Second, inflation expectations are unobservable and have to be approximated via an appropriate method. A commonly used method in the literature (e.g., Galí and Gertler 1999) is to approximate the unobserved inflation expectations in (8) by the corresponding realized future inflation, i.e., $E_t\pi_{t+1} = \pi_{t+1} - e_{t+1}$, where e_{t+1} denotes the rational prediction error. This approach, however, induces an extra disturbance to the underlying model which is likely to affect the accuracy of the estimation of the variance of the error term associated with the NKPC model. It also leads to econometric difficulties in the serial correlation test for the residuals in empirical estimations. To avoid these problems, we follow Pagan (1984) and approximate inflation expectations by projecting realized future inflation on the instrumental variable (IV) set

Z used in our empirical estimations (i.e., $E_t\pi_{t+1} = P_Z\pi_{t+1}$, where $P_Z = Z(Z'Z)^{-1}Z'$ is the projection matrix in terms of the IV set). It follows that this procedure will yield precisely the same coefficient estimates as those obtained by the IV estimation with the rational expectations approximation, while the standard errors will be different since the former ignores uncertainty in the estimation of the projection matrix (Pagan 1984).

Third, the real domestic output gap in the baseline estimations is obtained from the Hodrick-Prescott (HP) filter on the corresponding real GDP series (with smoothing parameter 1,600 for quarterly data). In robustness assessments, we also use deterministic linearly and quadratically detrended log real GDP to approximate the real output gap in model (8).

Fourth, the real foreign output gap is calculated as a weighted sum of the real GDP gap measure of the top ten U.S. major trading partners.[2] The weight for each trade partner in each year is determined by the percentage of the partner's total trade (exports and imports) to the United States over the total trade of the United States with the ten partners in that year. Table 1 presents the corresponding statistics for the trade weights of each of the ten countries to the United States over 1980–2012. It shows that the trade weights of different countries to the United States change over time. For example, the trade percentage of China to the United States witnessed a dramatic jump in the mid-1990s, rising from less than 5 percent before 1992 to above 7 percent in the late 1990s and around 20 percent over the most recent five years. Interestingly, however, the trade percentage of Japan to the United States has experienced a steady decline from the 1990s to the 2000s. After 2002, China took over Japan and became the second largest U.S. trading partner. A similar pattern of declining trade percentage to the United States can also be observed for Canada, which was the largest U.S. trading partner in most of the times.

These variations in the trade percentages are fully accounted for in the calculation of the foreign real GDP gap, i.e., $\hat{y}_t^* = \sum_{j=1}^{18} w_{j,t}\hat{y}_{j,t}$, where $w_{j,t}$ denotes the defined weight (i.e., trade

[2]The top ten major U.S. trading partners include Canada, China, France, Germany, Italy, Japan, Korea, Mexico, Netherlands, and the United Kingdom. The trade data and real GDP series for the ten countries were obtained from the International Monetary Fund.

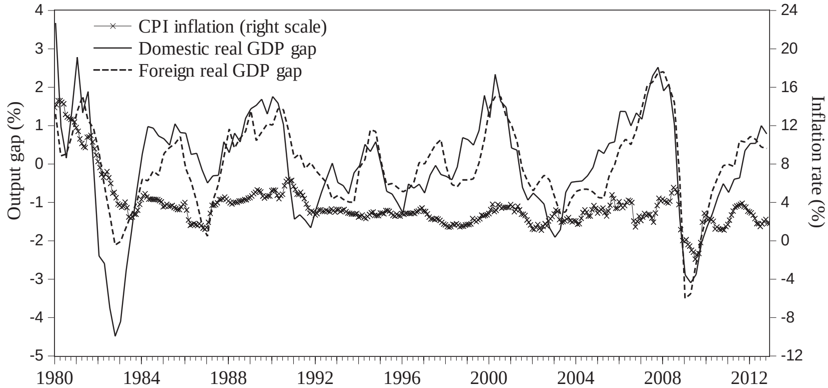
Table 1. Trade Percentage of U.S. Top Ten Major Trade Partners (%)

Year	CAN	CHN	FRA	GER	ITA	JAP	KOR	MEX	NETH	UK
1980	30.6	1.9	5.2	9.2	4.0	21.2	3.6	11.1	4.2	9.1
1981	30.9	2.0	4.8	7.9	3.9	22.1	3.8	11.4	4.0	9.2
1982	30.2	2.0	4.9	8.2	3.9	22.9	4.3	10.4	4.2	9.1
1983	32.5	1.7	4.4	7.9	3.5	23.5	4.9	9.4	3.9	8.4
1984	33.0	1.9	4.2	7.8	3.8	24.4	4.7	8.8	3.5	7.9
1985	31.6	2.2	4.4	8.2	4.1	25.7	4.5	8.9	3.2	7.3
1986	28.9	2.1	4.5	9.3	4.1	28.5	5.0	7.6	3.1	7.0
1987	29.9	2.4	4.4	9.0	3.9	26.4	5.9	8.0	2.8	7.3
1988	29.8	2.8	4.5	8.2	3.7	25.7	6.4	8.7	2.9	7.3
1989	30.0	3.4	4.5	7.6	3.6	25.4	6.1	9.4	3.0	7.1
1990	29.9	3.6	4.6	8.1	3.6	24.0	5.7	10.0	3.1	7.5
1991	29.6	4.4	4.8	8.0	3.5	23.7	5.5	10.8	3.1	6.8
1992	29.6	5.4	4.6	7.9	3.3	22.8	4.9	11.8	3.0	6.7
1993	30.8	6.1	4.2	7.0	2.9	22.8	4.7	11.9	2.7	7.0
1994	31.1	6.4	3.9	6.6	2.9	22.3	4.9	12.8	2.5	6.7
1995	31.0	6.8	3.6	6.8	3.0	21.7	5.7	12.3	2.6	6.4
1996	31.3	7.1	3.6	6.8	3.0	19.9	5.3	14.0	2.5	6.5
1997	31.2	7.6	3.7	6.7	2.8	18.4	4.7	15.4	2.7	6.8
1998	30.9	8.3	4.0	7.3	2.9	17.0	3.9	16.3	2.5	7.0
1999	30.9	8.5	3.9	7.1	2.8	16.4	4.6	16.7	2.4	6.7
2000	30.0	9.1	3.8	6.6	2.8	15.9	5.1	18.1	2.4	6.3
2001	29.9	10.0	4.0	7.1	2.7	14.6	4.6	18.3	2.3	6.5
2002	29.3	12.2	3.8	7.1	2.8	13.8	4.7	18.3	2.2	5.9
2003	29.3	14.2	3.5	7.3	2.8	12.8	4.6	17.5	2.4	5.7
2004	28.9	15.8	3.5	7.1	2.6	12.1	4.8	17.3	2.4	5.4
2005	29.2	17.5	3.3	7.0	2.6	11.5	4.3	17.0	2.4	5.3
2006	28.1	18.8	3.3	6.9	2.4	11.1	4.2	17.5	2.6	5.2
2007	27.7	19.8	3.4	7.2	2.5	10.4	4.1	17.1	2.6	5.3
2008	27.9	19.9	3.5	7.2	2.5	9.8	3.9	17.2	2.9	5.3
2009	25.5	22.3	3.6	6.8	2.3	8.8	4.1	18.1	2.9	5.5
2010	25.6	23.0	3.2	6.4	2.1	8.9	4.3	19.1	2.6	4.8
2011	25.8	22.4	3.0	6.4	2.2	8.5	4.4	19.9	2.9	4.6
2012	25.2	22.6	3.0	6.5	2.2	9.0	4.2	20.2	2.6	4.5

Note: The statistic is calculated as the ratio of the U.S. trade to each country as a percentage of total trade of the United States to all countries listed in the table; initial letters of each country's name are used as an acronym to represent the country.

percentage) at time t (quarterly observations within one year use the same weight of the year) and $\hat{y}_{j,t}$ is the real output gap measure for country j. Figure 2 plots the resulting foreign output gap series in conjunction with the U.S. real GDP gap and inflation series. It

Figure 2. U.S. Inflation, Domestic and Foreign Output Gaps: 1980:Q1–2012:Q4



Source: The author's calculations.

shows that the cyclical behavior of the domestic and foreign output gaps is similar in general but differs in details. In particular, the domestic GDP gap is more volatile than the foreign GDP gap before the late 1990s and vice versa during the most recent decade. This difference is reflected in the comparison of the dynamic evolution between domestic CPI inflation and the two output gaps: U.S. inflation appears to move more closely with the domestic output gap than with the foreign output gap before the late 1990s, and the scenario reverses afterwards. Whether this difference envisions structural changes in inflation dynamics model remains an empirical issue.

It may be noted that some studies (e.g., Martínez-García 2015) have shown that data quality and measurement issues also play an important role and may explain the lack of consensus as regards the effect of globalization on inflation in the literature. However, it should be noted that the relevant findings are different even if we examine the comparable results with similar data and measurement in the literature. Therefore, there is still room for model specification and structural break issues in accounting for the dispute on the nexus between globalization and inflation.

4. Empirical Results

4.1 Econometric Issues

As already noted in section 1 of this paper, the U.S. economic globalization level has increased substantially since the 1990s, and this change may shift the mechanism of the impact of globalization on the inflation dynamics. In particular, the empirical sample in our analysis covers a relatively long period from 1980 to 2012, which witnesses profound changes in the U.S. integration to the world economy and its macroeconomic dynamics. While the link between economic globalization and inflation dynamics makes it plausible that such changes may lead to structural breaks in the parameters of the NKPC model (8), any such effect and its timing depend on the behavior of economic agents. Since the dates of potential change points are therefore unknown, we perform break tests using the methodology proposed by Andrews (1993).

Prior to examining the structural break tests, several econometric issues in estimating the baseline model should be noted. First, inflation expectations in model (8) may be influenced by information relating to the current period. In addition, the real variables are also likely to be correlated with the contemporaneous noise, since demand shocks may influence both variables. Therefore, we use IV, or more generally the generalized method of moments (GMM) estimator to estimate model (8) and pin down the endogeneity problem.

The baseline IV set used in estimating (8) consists of two lags of each of the domestic and foreign output gaps, effective federal funds rate, and M2 growth rate, plus the lags of inflation included in the model (and a constant). The optimal lag order in the NKPC model (8) is specified by Akaike information criteria (AIC) and Godfrey's (1994) IV serial correlation test. In addition, the baseline estimations are verified through Hansen's (1982) J-test for over-identifying restrictions and the Stock and Yogo (2003) generalized F-test for weak IV.

Note that the Godfrey IV serial correlation test is implemented by adding appropriate lagged residuals from the initial estimation to the regressors from the initial model and checking their joint significance by the Lagrange multiplier (LM) principle. This test is

Table 2. Results of Andrews Unknown Break Point Tests for Model 8

	$\boldsymbol{all}$	$\boldsymbol{c}$	$\boldsymbol{\gamma_e}$	$\boldsymbol{\gamma_b}$	$\boldsymbol{\alpha_i}$	$\boldsymbol{\delta_d}$	$\boldsymbol{\delta_f}$
p-value	0.000	0.046	0.020	0.001	0.085	0.057	0.060
Break Date	2005:Q4	1987:Q1	1992:Q3	1991:Q3	1986:Q2	1990:Q4	1993:Q1

Notes: The estimated equation is given by model 8 with a sample spanning 1980:Q1–2012:Q4 prior to lag adjustment. Optimal autoregressive lag order in the NKPC model is specified by AIC and IV serial correlation test (with maximum eight lags). Inflation expectations are measured by fitted values of regressing the realized future inflation on the baseline IV set. The baseline IV set includes two lags of each of the domestic and foreign output gaps, effective federal funds rate, and M2 growth rate, plus the lags of inflation included in the model (and a constant).The structural break tests are implemented over central 70 percent of the underlying sample; break date corresponds to the break point at which the maximum test statistic is achieved.

used to check the possibility of serial correlation in the IV estimations with null hypothesis of no serial correlation. Therefore, a large p-value indicates no significant serial correlation in the regression and vice versa. The Stock-Yogo weak instrument test provides diagnostic information on to what extent the underlying instruments are weak in the estimation. The statistics reported are the Cragg-Donald statistics, with larger values indicating stronger IV sets.

Based on the preceding design, we carry out formal unknown structural break tests. Specifically, we employ the supreme likelihood ratio (LR) test of Andrews (1993) to test for unknown structural breaks in model (8). The test is designed to test for the null hypothesis of no structural break in the underlying parameters of interest. The corresponding p-values of the tests are computed using the method of Hansen (1997). We set a conventional searching interval of the central 70 percent of the full sample to allow a minimum of 15 percent of effective observations contained in both pre- and post-break periods to avoid extreme statistic results.

4.2 Baseline Results

Table 2 summarizes the results of the Andrews (1993) unknown structural break tests for model (8). The break tests are performed on all the coefficients overall and then on the individual coefficients.

Specifically, the first row in table 2 provides notational information for the coefficients in the break tests, with the first statistic (denoted *all*) testing for stability of all the coefficients in (8), while the other results refer to individual tests for the indicated coefficients. The second row reports p-values associated with the corresponding break test statistics for the null hypothesis of no structural change, while the third row, labeled *break date*, represents the estimated break date corresponding to the *Sup-LR* statistic.

As can be seen from the results in table 2, p-values for all break test statistics are quite small (significant at different significance levels), indicating structural instability of the underlying model. The break date statistics do not provide a uniform break point. Nonetheless, the structural break for the key coefficients of our interest, namely the coefficients on the domestic and foreign output gaps (δ_d *and* δ_f), as well as the forward- and backward-looking coefficients concentrates in the early 1990s. The break dates for the overall coefficients (*all*), the intercept (c), and the autoregressive coefficients (α_i) are close to the polar points in the searching interval and hence may be unreliable.

Since the structural break tests with unknown points fail to identify a clear break point for the underlying model, we need to produce additional empirical evidence based on a more comprehensive set of formal parameter constancy tests in support of the identified break points in the early 1990s. Therefore, we define dummy variables based on change points in 1990:Q4, 1991:Q3, 1992:Q3, and 1993:Q1 to investigate the nature of these structural breaks. Since our focus is the structural break in output gap measures and also to alleviate the associated collinearity induced by the multiplicative dummy variables, we use the following model to carry out dummy tests for each individual break point, viz.,

$$\begin{aligned}\pi_t = c + \gamma_e E_t \pi_{t+1} + \gamma_b \pi_{t-1} + \sum_{i=1}^{m-1} \alpha_i \Delta\pi_{t-i} + \delta_d \hat{y}_t + \delta_f \hat{y}_t^* \\ + \delta_d' \hat{y}_t d_t + \delta_f' \hat{y}_t^* d_t + u_t, \end{aligned} \tag{9}$$

where d_t denotes the dummy variable with zero before the break point and unity otherwise.

Table 3. Results of Dummy Tests for Model 9

Break Point	1990:Q4	1991:Q3	1992:Q3	1993:Q1
δ'_d	-1.04^{**}	-1.66^{**}	-1.35	-1.647
	(0.433)	(0.718)	(0.927)	(1.205)
δ'_f	0.973^{**}	1.430	1.038	1.366
	(0.459)	(0.940)	(0.816)	(1.245)
Joint Test	0.059^{*}	0.021^{**}	0.344	0.310

Notes: The estimated equation is given by model 9 over the sample period from 1980:Q1 to 2012:Q4 prior to lag adjustment. The Bartlett kernel with Newey-West (fixed bandwidth) HAC-robust standard errors are reported in parentheses. "Joint Test" denotes the p-value for a joint significance test (F-statistic) for the coefficients on all dummy variables; * and ** denote statistical significance at the 10 and 5 percent levels, respectively.

The results for the dummy tests with the four possible break points are reported in table 3 using the sample 1980:Q1–2012:Q4. Conditional on the break point of 1990:Q4, both individual coefficient tests and the joint significance test are statistically significant, which provides clear evidence that the coefficients on both domestic and foreign output gap measures experience significant structural change in 1990:Q4. As a comparison, the joint test and tests for coefficients on the domestic output gap are significant (but the coefficient on the foreign output gap is not significant) for the break point of 1991:Q3. The similar results for the first two break points are unsurprising since these two points are very close to each other. For the break points of 1992:Q3 and 1993:Q1, the results do not provide clear evidence of a significant change in coefficients (all the dummy tests are insignificant).

Overall, the results of the structural break tests confirm that the inflation dynamics model (8) for the United States experiences a significant structural break in the early 1990s, and the evidence of change relates to the domestic and foreign output gaps as well as forward- and backward-looking components in the NKPC model. We have now obtained a structural break point based on which we can investigate the nature of changes in the United States' inflation dynamics and compare the impact of globalization on inflation dynamics over different sample periods when the break in the underlying coefficients is recognized at the early 1990s. In what follows, we will take the break date pertaining to the coefficient on the output gap (i.e., 1990:Q4) as the benchmark time to split the sample. The slightly different break points of 1991, 1992, and 1993 reported in table 2 will be examined in the robustness analysis. We will also evaluate the changing pattern of the coefficient estimates on the domestic and foreign output gap over recursive and rolling samples in the robustness analysis.

Table 4 reports GMM estimates of model (8) over the whole sample and pre- and post-1990 periods for the forward- and backward-looking inflation coefficients and the domestic and foreign output gap measures, in conjunction with relevant diagnostic statistics. The diagnostic test statistics in table 4 indicate that the specification of model (8) is generally free from significant serial correlation and the IV choice is valid and relatively strong in most cases. The p-values of the joint significance tests on the extra lagged inflation (from order two onwards) are smaller than 0.01 in all regressions, indicating a statistically significant role of the extra lagged inflation in the empirical NKPC model.

More importantly, the baseline results reported in panels A, B, and C in table 4 reveal significant changes in the impact of the domestic and foreign output gaps on U.S. inflation over different sample periods. Specifically, panel A shows that if the structural break is neglected, the domestic output gap drives inflation significantly with the coefficient estimate of 0.105. In contrast, the coefficient estimate on the foreign output gap is very small (−0.004) and statistically insignificant. Albeit the magnitude of the coefficient estimate on the domestic output gap declines when the convex restriction $\gamma_e + \gamma_b = 1$ is imposed, it is still much higher than the coefficient estimate on the foreign output gap and remains statistically significant.

When the structural break is accounted for, however, the results changed distinctively. Panel B and panel C provide evidence that the impact of the domestic and foreign output gaps on inflation has changed (switched) significantly before and after 1990. The coefficient estimate of the domestic output gap falls substantially from a large and significant value (0.166) pre-1990 to a small and insignificant value (0.001) post-1990. Conversely, the coefficient estimate of the foreign output gap has risen from an insignificant and small

Table 4. (GMM) Estimation Results of the Inflation Dynamics for the United States

Sample	**Baseline Estimates**				**Diagnostic Tests**			
	γ_e	γ_b	δ_d	δ_f	$p(\alpha_i)$	p-auto	p-J	WeakIV
A. 1980:Q1–2012:Q4								
	0.394***	0.567***	0.105***	−0.004	0.000	0.079	0.855	14.4###
	(0.018)	(0.07)	(0.011)	(0.005)				
$\gamma_e + \gamma_b = 1$	0.491***	0.509***	0.057***	0.019	0.000	0.034	0.884	14.4###
	(0.021)	(0.021)	(0.021)	(0.029)				
B. Pre-1990:Q4								
	0.320	0.630***	0.166*	−0.052	0.000	0.410	0.616	4.02
	(0.413)	(0.213)	(0.089)	(0.287)				
$\gamma_e + \gamma_b = 1$	0.444***	0.556***	0.128	−0.063	0.000	0.054	0.678	4.02
	(0.148)	(0.148)	(0.136)	(0.201)				
C. Post-1990:Q4								
	0.520***	0.454***	0.001	0.080*	0.000	0.057	0.587	6.33##
	(0.047)	(0.040)	(0.014)	(0.042)				
$\gamma_e + \gamma_b = 1$	0.539**	0.461***	−0.013	0.085**	0.000	0.003	0.846	6.33##
	(0.048)	(0.048)	(0.008)	(0.039)				

Notes: The estimated equation is given by model 8. Inflation expectations are measured by fitted values of regressing the realized future inflation on the baseline IV set (IV is the same as in table 2). The Bartlett kernel with Newey-West (fixed bandwidth) HAC-robust standard errors are reported in parentheses. $p(\alpha_i)$ is the p-value of a joint significance test on lagged inflation beyond order one; p-auto, p-J, and weakIV refer to p-values of Godfrey's (1994) IV serial correlation test, Hansen's (1982) J-test, and Stock and Yogo's (2003) weak instrumental variables test (critical values for the weak IV test are provided in Stock and Yogo 2003), with ####, ###, ##, and # denoting statistically signficantly strong IV (at the 5 percent significance level) when the desired maximal bias of the IV estimator relative to OLS is specified to be 5, 10, 20, and 30 percent, respectively. *, **, and *** denote statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

negative value (−0.05) to a significant and positive value (0.08). A similar scenario can be observed when the convex restriction is imposed.

The results in table 4 suggest that there is a significant decline in the sensitivity of U.S. inflation to the domestic output gap since the early 1990s. In addition, the impact of globalization, as captured by the foreign output gap, on domestic inflation has increased dramatically since 1990. The comparison between the coefficient estimates on the domestic and foreign output gaps clearly shows that the foreign output gap plays a more prominent role than the domestic counterpart in the domestic inflation process after 1990.

To summarize, the coefficient estimates of the domestic and foreign output gap measures provide our main finding from table 4, namely that the impact of the domestic and foreign output gaps on U.S. inflation has changed significantly. The foreign output gap plays a more important role than the domestic output gap in affecting the domestic inflation process over the most recent two decades. This finding indicates that inflation dynamics in the United States has shifted significantly since the early 1990s and the impact of globalization via the foreign excess demand has indeed risen accordingly. The next section assesses the robustness of this finding.

4.3 Robustness Assessments

To assess the robustness of the baseline finding that the impact of globalization on inflation dynamics has changed, we carry out three sets of sensitivity exercises. First, we investigate whether the finding is robust to slightly different break points of 1991, 1992, and 1993, which are shown in table 2. Second, we assess the robustness of the baseline finding by dropping the domestic and foreign output gap from the NKPC model. If the foreign economic slack does not exert any influence on the sensitivity of inflation to the domestic output gap, the coefficient estimates on the domestic output gap in models with and without foreign output gap should be similar. Third, we examine recursive estimates of the coefficients on the domestic and foreign output gaps to provide further evidence on the changing pattern of the impact of globalization on inflation in the United States.

Table 5 reports the results for the first exercise. Panel A shows that the domestic output gap is significant for the pre-1991 period but insignificant (and small) for the post-1991 era, while the converse applies to the foreign output gap. The results for break points being 1992 and 1993 in panels B and C provide a similar scenario as the baseline finding. Although the coefficients on the foreign output gap are statistically insignificant over the post-break periods (which may be unsurprising because the underlying variables in the model have been very smooth during these periods), the comparison between the coefficient estimates on the domestic and foreign output gaps clearly shows that the foreign output gap plays a more prominent role than the domestic counterpart in the domestic inflation process after the early 1990s.

However, obtaining statistical significance for the foreign output gap over the post-1990 period is difficult. Conditioning on the alternative break points after 1990:Q4, the coefficient of the global output gap is no longer statistically significant. Based on these results, one might argue that the empirical evidence in support of the globalization hypothesis is weak from the perspective of statistical significance. The statistical insignificance, however, may reflect the fact that the break point for the foreign output gap is different from those for other variables in the model. In addition, although the statistical significance for the foreign output gap is not clear from table 5, the following additional exercises provide further evidence and confirm the changing role of the domestic and foreign output gap in driving domestic inflation.

Table 6 summarizes the results for the second exercise. Panel A and panel B provide the estimation results when either the domestic or foreign output gap is dropped from the NKPC model. It appears to be difficult to distinguish the roles of the domestic and foreign output gaps in affecting the inflation process before 1990 in terms of the magnitude of the coefficient estimates, but statistically speaking, the domestic output gap is significant while the foreign output gap is insignificant. For the post-1990 period, it is clear that the domestic output gap plays a small and insignificant role but the foreign output gap plays a large and significant role in driving domestic inflation. These results reinforce the conclusion that the foreign economic slack plays a more important role than the domestic output in affecting U.S. inflation after the early 1990s.

Table 5. Robustness: Alternative Break Dates

	Baseline Estimates				Diagnostic Tests			
	γ_e	γ_b	δ_d	δ_f	$p(\alpha_i)$	p-auto	p-J	WeakIV
A. Break = 1991:Q3								
Pre-1991	0.323	0.631***	0.166*	−0.058	0.000	0.433	0.605	4.35
	(0.189)	(0.120)	(0.096)	(0.157)				
Post-1991	0.519***	0.466***	−0.041	0.121	0.000	0.057	0.693	6.72##
	(0.068)	(0.102)	(0.088)	(0.172)				
B. Break = 1992:Q3								
Pre-1992	0.306***	0.639***	0.165***	−0.060	0.000	0.469	0.566	4.54#
	(0.026)	(0.025)	(0.029)	(0.058)				
Post-1992	0.532***	0.462***	−0.051	0.129	0.000	0.059	0.681	7.55##
	(0.115)	(0.088)	(0.197)	(0.375)				
C. Break = 1993:Q1								
Pre-1993	0.307***	0.640***	0.165*	−0.063	0.000	0.516	0.542	3.85
	(0.161)	(0.109)	(0.094)	(0.167)				
Post-1993	0.543***	0.466***	−0.055	0.126	0.000	0.055	0.705	8.09##
	(0.062)	(0.084)	(0.074)	(0.174)				

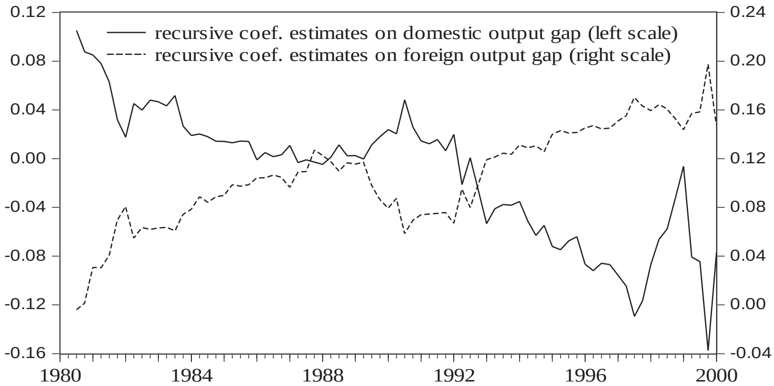
Notes: See table 4.

Table 6. Robustness: Dropping Domestic or Foreign Real Variables

Sample	Baseline Estimates				Diagnostic Tests			
	γ_e	γ_b	δ_d	δ_f	$p(\alpha_i)$	p-auto	p-J	WeakIV
A. Pre-1990								
No Foreign	0.290	0.646***	0.144*		0.000	0.367	0.565	5.05#
	(0.247)	(0.162)	(0.084)					
No Domestic	0.504**	0.477***		0.154	0.000	0.133	0.671	4.54
	(0.204)	(0.108)		(0.225)				
B. Post-1990								
No Foreign	0.537***	0.454**	0.070		0.000	0.091	0.671	17.47####
	(0.118)	(0.184)	(0.096)					
No Domestic	0.394***	0.463***		0.106**	0.000	0.028	0.840	16.62####
	(0.036)	(0.100)		(0.063)				

Notes: "No Foreign" means that the regression excludes the foreign output gap, and "No Domestic" is defined analogously; other notations follow table 5.

Figure 3. Recursive Estimates of δ_d and δ_f

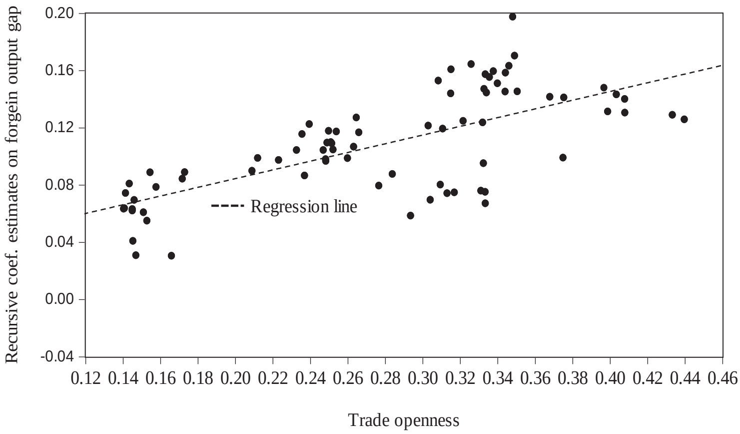


Note: Recursive samples start from 1980:Q1 to 2000:Q1 with the end period 2012:Q4 being fixed.

The design of the third exercise provides another useful check on the robustness of the baseline finding over recursive samples. By construction, the recursive estimations produce coefficient estimates on the domestic and foreign output gaps over the recursive samples starting from 1980 onwards with the end period (2012:Q4) fixed. By looking at the recursive estimates, we can (presumably) evaluate the changing pattern of the coefficient estimates on the domestic and foreign output gaps. Figure 3 plots the results of the recursive estimations. It shows that the coefficients on the domestic output gap generally manifest a declining trend while the coefficients on the foreign output gap grow over the forward recursive samples. This result is consistent with our finding of the increasingly important role of the foreign economic slack in affecting domestic inflation. Although the recursive estimates might not be very accurate by nature, the results corroborate the finding of a great deal of the baseline empirical work.

Based on the recursive estimates, we further obtain the scatter plot of the coefficient estimates for the foreign output gap against globalization (i.e., trade openness) over time (1980:Q1–2001:Q1) in figure 4. The upward-sloping regression line fitted in the scatter plot suggests that the coefficient estimates on the foreign output gap in general move in the same direction as the level of U.S. economic

Figure 4. Scatter Plot of Recursive Estimates of δ_f against Trade Openness (1980:Q1–2001:Q1)



globalization. This result indicates that the impact of the foreign output gap on inflation increases when the level of U.S. economic globalization rises.

5. Discussion

The empirical results in section 4 provide evidence of a significant structural break in the NKPC model for the United States. The results also indicate that the roles that domestic and foreign economic slack played in U.S. inflation have switched in the early 1990s: the foreign output gap exerts larger impact than the domestic output gap on inflation since the early 1990s. Omitting such a structural break, as the results pertaining to the whole sample estimation indicate, can lead to a fragmented conclusion that the domestic output gap plays a predominant role in the inflation process during the entire period of the past three decades.

The important role of the foreign output gap in the NKPC model since the 1990s raises a realistic concern that the problems facing the Federal Reserve are likely to be complicated by the rising economic globalization of the U.S. economy. More specifically, because the Phillips curve is an indispensable component in monetary policy analysis, it is natural that the impact of the globalization factor on inflation can also transmit to other macroeconomic variables in policy analysis frameworks. Suppose we analyze the issue in a standard three-equation model as in Woodford (2009) with an IS curve, a Phillips curve, and a policy reaction function. The impact of globalization on inflation will transmit to domestic real output through the IS curve and to monetary policy via the policy reaction function.

Interestingly, however, Woodford (2009) carries out a formal theoretical analysis on the possible impact of globalization on this traditional monetary policy transmission process, and his simulation results appear to suggest that increased globalization engenders no substantial reduction of the effects of domestic monetary policy on the domestic economy. Through the theoretical designs, Woodford provides a comprehensive and valuable discussion on a wide range of ways that globalization might weaken the central bank's ability to influence the economy.

Woodford carefully interprets in his conclusion that his results mainly suggest that increased globalization should not eliminate the influence of domestic monetary policy over domestic inflation, but do not mean that the degree of openness of an economy is of no significance for the conduct of monetary policy. Indeed, increased international trade in financial assets, consumer goods, and factors of production should lead to quantitative changes in the magnitudes of various key response elasticity relevant to the transmission mechanism for monetary policy (Woodford 2009). In particular, changes in the degree of goods market integration affect the quantitative specification of both the aggregate demand and aggregate supply blocks in Woodford's analysis.

Our work also provides empirical results that are complementary to the comprehensive studies of Mishkin (2007, 2009) which articulate that the decline in the sensitivity of inflation to the domestic output gap in the United States is the result of better monetary policy that has anchored inflation expectations. Our results do not exclude the potential contribution of better monetary policy on the flattening of the Phillips curve. The finding does not mean that the rising globalization of the U.S. economy will eliminate the Federal Reserve's capacity to stabilize domestic inflation, nor do we regard the rising globalization as a fatal fear to the national economy. The bottom line, however, is that the central bank can gear up its capacity to control inflation by appropriate coordination with other central banks. Even without material coordinated actions, the central bank may still be able to increase the precision of its relevant forecasts and thereby improve the effect of its policies on domestic economy by adding the global development into its forecasting information set or augmenting its policy analysis framework with globalization factors.

On a minimum (and positive) side, the baseline finding in the present paper calls attention to the rising globalization that may engender material forces for central bankers to confront more practical issues than the traditional issues in a closed economy. The changing degree of globalization also makes the issue of change over time in the correct quantitative specification of the models used in a central bank a more pressing one to consider (Woodford 2009).

6. Conclusions

The past three decades have been characterized by a steady progress of global economic integration. The U.S. economy has also become increasingly more open. This progress of globalization may have led to important changes in U.S. inflation dynamics. This paper constructed an extended New Keynesian Phillips-curve model from microeconomic foundations and showed that globalization factors (i.e., foreign output gap) can be incorporated into such a model. We proposed that this model may have experienced structural breaks given the fact that the level of globalization of the U.S. economy has changed substantively over the period of 1980 to 2012.

Our empirical investigations showed that there is a significant structural change in the early 1990s in the extended NKPC model for the United States, after which inflation responds much more to the foreign economic slack than to the domestic output gap. This finding indicates that the Federal Reserve should take into account the developments in global economic performance in the monetary policymaking process. It further implies that while the higher level of globalization may help subdue and stabilize domestic inflation by, for example, stabilizing global economic slack during globally tranquil times, a negative global economic environment can also present extra challenges for the domestic policymakers. Therefore, studies that neglect the role of foreign economic slack are likely to underestimate the impact of globalization factors on domestic inflation dynamics, at least for the period post-1990.

Appendix

This appendix provides details in developing a dynamic, micro-founded inflation dynamics model for an open economy with sticky prices. The model is an extension of recent developments in the open-economy dynamic stochastic general equilibrium literature. To facilitate notations, we will use variables with an i-index to refer to economy i and variables without an i-index to refer to the open economy being modeled. Variables with subscripts D and F refer to domestic variables and foreign variables, respectively, while variables with a star superscript refer to the world economy as a whole.

Households

The objective of a representative household in a typical open economy is to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t U(C_t, N_t), \tag{10}$$

where N_t denotes labor services provided by the household, and C_t is a composite consumption index composed of consumptions of both domestic goods and imported goods, viz.

$$C_t \equiv \left[(1-a)^{\frac{1}{\eta}} (C_{D,t})^{\frac{\eta-1}{\eta}} + \alpha^{\frac{1}{\eta}} (C_{F,t})^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}, \tag{11}$$

where $C_{D,t}$ and $C_{F,t}$ are indexes for consumptions of domestic and imported goods, respectively, both of which are consumed by domestic households and given by the continuum version of the standard constant elasticity of substitution (CES) function. Parameter α is related to the degree of home bias in preferences, and parameter η measures the substitutability between domestic and foreign goods, from the perspective of the domestic consumers.

The maximization of (10) is subject to a sequence of budget constraints of the form

$$\int_0^1 P_{D,t}(j)C_{D,t}(j)dj + \int_0^1 \int_0^1 P_{i,t}(j)C_{i,t}(j)djdi + E_t(Q_{t,t+1}Z_{t+1}) \\ \leq Z_t + W_tN_t + T_t, \quad (12)$$

where $P_{D,t}(j)$ is the price of variety j produced domestically and $P_{i,t}(j)$ is the price of variety j imported from country i (expressed in domestic currency). Z_{t+1} is the nominal payoff in period $t+1$ of the portfolio held at the end of period t, W_t denotes the nominal wage, T_t denotes lump-sum taxes, and $Q_{t,t+1}$ is the stochastic discount factor for one-period-ahead nominal payoffs relevant to the domestic household.

In the open economy, the households' demand functions can be categorized as

$$C_{D,t}(j) = \left(\frac{P_{D,t}(j)}{P_{D,t}}\right)^{-\varepsilon} C_{D,t}, C_{i,t}(j) = \left(\frac{P_{i,t}(j)}{P_{i,t}}\right)^{-\varepsilon} C_{i,t}, \quad (13)$$

where $P_{D,t}(j) \equiv (\int_0^1 P_{D,t}(j)^{1-\varepsilon}dj)^{\frac{1}{1-\varepsilon}}$ is the price index for goods produced domestically, $P_{i,t}(j) \equiv (\int_0^1 P_{i,t}(j)^{1-\varepsilon}dj)^{\frac{1}{1-\varepsilon}}$ is a price index for goods imported from country i, and ε is the elasticity of substitution between varieties.

Let $P_{F,t} \equiv (\int_0^1 P_{i,t}^{1-\gamma}di)^{\frac{1}{1-\gamma}}$ be the price index for imported goods (where γ measures the substitutability between goods produced in different foreign countries); the optimal allocation of consumptions on imported goods by country of origin implies

$$C_{i,t} = \left(\frac{P_{i,t}}{P_{F,t}}\right)^{-\gamma} C_{F,t}. \quad (14)$$

Since domestic households' consumptions are composed of domestically produced goods and imported goods, the overall domestic price index (DPI) can be written as $P_t \equiv [(1-\alpha)(P_{D,t})^{1-\eta} + \alpha(P_{F,t})^{1-\eta}]^{\frac{1}{1-\eta}}$. As such, the optimal allocation of consumptions between domestic and imported goods is given by

$$C_{D,t} = (1-\alpha)\left(\frac{P_{D,t}}{P_t}\right)^{-\eta} C_t, \quad C_{F,t} = \alpha\left(\frac{P_{F,t}}{P_t}\right)^{-\eta} C_t. \quad (15)$$

Accordingly, total consumption by domestic households is given by $P_{D,t}C_{D,t} + P_{F,t}C_{F,t} = P_tC_t$. Thus the period budget constraint can be rewritten as

$$P_tC_t + E_t(Q_{t,t+1}Z_{t+1}) \leq Z_t + W_tN_t + T_t. \tag{16}$$

In addition, we assume the period utility function takes the form $U(C, N) \equiv (1-\sigma)^{-1}C^{1-\sigma} - (1+\varphi)^{-1}N^{1+\varphi}$. The remaining optimality conditions for the households' problem can be written as

$$C_t^{\sigma}N_t^{\varphi} = W_t/P_t \tag{17}$$

and

$$\beta\left(\frac{C_t}{C_{t+1}}\right)^{\sigma}\left(\frac{P_t}{P_{t+1}}\right) = Q_{t,t+1}. \tag{18}$$

Taking conditional expectations on both sides of (18), we obtain a stochastic Euler equation

$$\beta R_t E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma}\left(\frac{P_t}{P_{t+1}}\right)\right] = 1, \tag{19}$$

where $R_t = (E_tQ_{t,t+1})^{-1}$ denotes the gross return on a risk-free one-period discount bond.

It is useful to note that (17) and (19) can be respectively written in log-linearized form as

$$\begin{cases} w_t - p_t = \sigma c_t + \varphi n_t \\ c_t = E_t\{c_{t+1}\} - \frac{1}{\sigma}(r_t - E_t\{\pi_{t+1}\} - \rho), \end{cases} \tag{20}$$

where lowercase letters denote the natural logarithm of the respective variables, $\rho \equiv \beta^{-1} - 1$ is the time discount rate, and $\pi_t \equiv p_t - p_{t-1}$ (with $p_t \equiv Ln(P_t)$) is the overall domestic price inflation.

In what follows we define the bilateral terms of trade between the domestic economy and country i as $S_{i,t} = P_{i,t}/P_{D,t}$. The effective terms of trade can be approximated by the log-linear expression $s_t = \int_0^1 s_{i,t}di$. Combining the definition of P_t in (15) with the purchasing power parity (PPP) condition, we can obtain the link between the DPI, $P_{D,t}$, and $P_{F,t}$, which can be further used to obtain the relationship between overall domestic price inflation π_t (defined as the rate of change in the index of overall domestic goods prices) and domestically produced goods price inflation $\pi_{D,t}$, viz.

$$\pi_t = \pi_{D,t} + \alpha \Delta s_t. \tag{21}$$

Let $e_t \equiv \int_0^1 e_{i,t} di$ denote the log nominal effective exchange rate and $p_t^* \equiv \int_0^1 P_{i,t}^i di$ denote the log world price index. Assuming that the law of one price holds at all times and combining the definition of $P_{F,t}$ in (14) with the definition of the terms of trade, we can obtain the linear expression for the terms of trade, the domestic price, and the world price as

$$s_t = e_t + p_t^* - p_{D,t}. \tag{22}$$

Let $q_t = \int_0^1 q_{i,t} di$ be the log effective real exchange rate (where $q_{i,t}$ is the log bilateral real exchange rate). It follows that

$$q_t = s_t + p_{D,t} - p_t = (1-\alpha)s_t. \tag{23}$$

Next, we consider the relationship between domestic consumption expenditures and the world consumption expenditures. As in Galí and Monacelli (2005), a first-order condition analogous to (18) also holds for the household in any other country i, viz.

$$\beta \left(\frac{C_{t+1}^i}{C_t^i} \right)^{-\sigma} \left(\frac{P_t^i}{P_{t+1}^i} \right) \left(\frac{\varpi_t^i}{\varpi_{t+1}^i} \right) = Q_{t,t+1}, \tag{24}$$

where ϖ denotes bilateral nominal exchange rate. Combining (18), (23), and (24), we obtain

$$c_t = c_t^* + \frac{1}{\sigma} q_t = c_t^* + \left(\frac{1-\alpha}{\sigma} \right) s_t, \tag{25}$$

where $c_t^* \equiv \int_0^1 c_t^i di$ is the log world consumption index.

Firms

We assume that a typical domestic firm produces differentiated goods with a linear technology represented by the production function of $Y_t(j) = A_t N_t(j)$, A_t denotes technology, and N_t is defined as in (10). Hence, the real marginal cost is given by

$$mc_t = -v + w_t - p_{D,t} - a_t, \tag{26}$$

where $v \equiv -\log(1-\tau)$ (τ denotes an employment subsidy) and $a_t \equiv \ln(A_t)$. Let y_t be the log domestic real output, analogously defined as the one introduced for consumption. It can be shown that

$$y_t = a_t + n_t, \tag{27}$$

where $n_t \equiv \log(N_t)$.

For the pricing behavior of the domestic firms, we assume an economic environment similar to Calvo's (1983) model, in which firms are able to revise their prices in any given period with a fixed probability $(1-\theta)$. In addition, we assume both "forward-" and "backward-looking" firms coexist in the economy with a proportion of ω and $(1-\omega)$, respectively. Further, we extend the rule of the recent pricing behavior of the backward-looking firms to incorporate a weighted process of past inflation, instead of stylized one lag of inflation inertia.

Specifically, based on the regular assumptions in Calvo's model and log-linear approximations, it is possible to obtain the (log) aggregate price level by

$$p_t = \theta p_{t-1} + (1-\theta)p_t^{new}, \tag{28}$$

where p_t^{new} is the new price set in period t. Let p_t^f be the price set by forward-looking firms and p_t^b be the price set by backward-looking firms at time t. The new price (relative to the aggregate price) can be expressed as a convex combination of p_t^f and p_t^b, viz.

$$p_t^{new} - p_t = \omega(p_t^f - p_t) + (1-\omega)(p_t^b - p_t). \tag{29}$$

Next, following Woodford (2003), the pricing behavior of the forward-looking firms can be written as

$$p_t^f - p_t = \theta\varsigma \sum_{s=0}^{\infty} (\theta\varsigma)^s E_t \pi_{t+s+1} + (1-\theta\varsigma) \sum_{s=0}^{\infty} (\theta\varsigma)^s E_t mc_{t+s}, \tag{30}$$

where ς denotes a subjective discount factor. Iterating (30) gives

$$p_t^f - p_t = \theta\varsigma E_t \pi_{t+1} + (1-\theta\varsigma) mc_t + \theta\varsigma E_t(p_{t+1}^f - p_{t+1}). \tag{31}$$

Assume a rule of thumb in the pricing setting, viz.

$$p_t^b - p_t = p_{t-1}^{new} - p_t + \pi_{t-1}. \tag{32}$$

As emphasized in the literature, this is an elegant innovation in that the backward-looking firms can now set their prices to the average price determined in the most recent price adjustments with a correction for inflation. However, inflation inertia in (32) is confined to one single lag, which may neglect the importance of other historical inflation in predicting current inflation. Therefore, it would be more reasonable to replace π_{t-1} in (32) with a weighted average of inflation over several periods in the past. This modification effectively mitigates a serial correlation problem in empirical analysis. As such, we extend (32) in the following process:

$$p_t^b - p_t = (p_{t-1}^{new} - p_t) + \pi_{t-1} + \rho^*(L)\Delta\pi_{t-1}, \tag{33}$$

where $\rho^*(L) = \rho_1^* + \rho_2^* L + \cdots + \rho_m^* L^{m-1}$ is a polynomial in lag operator.

Combining (28)–(33), it can be shown that the dynamics of domestic inflation in terms of real marginal cost are described by an equation analogous to the one associated with a closed economy, viz.

$$\pi_t = \gamma_e E_t \pi_{t+1} + \gamma_b \pi_{t-1} + \sum_{i=1}^{m-1} \alpha_i \Delta\pi_{t-i} + \lambda mc_t, \tag{34}$$

where coefficients of (34) are functions of the deeper parameters in (28)–(33).

Equilibrium and Inflation Dynamics Model

Combining goods market clearing conditions in the open economy with the definition of aggregate domestic output, we can obtain the following expression:

$$y_t = c_t + \alpha\gamma s_t + \alpha(\eta - \sigma^{-1})q_t = c_t + \frac{\alpha\varpi}{\sigma} s_t, \tag{35}$$

where $\varpi \equiv \sigma\gamma + (1-\alpha)(\sigma\eta - 1)$. Additionally, combining the world goods market clearing conditions with (18), (24), and (25), we can derive the relationship between real domestic output and real world output as

$$y_t = y_t^* + \frac{1}{\sigma_\alpha} s_t, \tag{36}$$

where $\sigma_\alpha \equiv [(1-\alpha)+\alpha\omega]^{-1}\sigma$.

Next, making use of (20), (25), (27), and (35), we can rewrite (26) as

$$mc_t = v + xi_d y_t + \xi^* y_t^* + u_t, \tag{37}$$

where $\xi_d \equiv (\sigma_\alpha + \varphi)$, $\xi^* \equiv (\sigma - \sigma_\alpha)$, and $u_t \equiv -(1+\varphi)a_t$. By definition of the real output gap (log-deviation from a steady-state level of output), the equation (37) can be rewritten as

$$mc_t = c^* + \xi_d \hat{y}_t + \xi^* \hat{y}_t^* + u_t, \tag{38}$$

where $c^* \equiv -v + \mu(\xi_d + \xi^*)$, and $\hat{y}_t$ and $\hat{y}_t^*$ are the real domestic output gap and the real world output gap, respectively.

Substituting (38) for mc_t in (34), we obtain an open-economy generalization of the extended New Keynesian Phillips curve (NKPC) that incorporates richer inflation dynamics and world excess demand, viz.

$$\pi_t = c + \gamma_e E_t \pi_{t+1} + \gamma_b \pi_{t-1} + \sum_{i=1}^{m-1} \alpha_i \Delta\pi_{t-i} + \delta_d \hat{y}_t + \delta_f \hat{y}_t^* + \lambda_t, \tag{39}$$

where $c \equiv \lambda c^*$, $\delta_d \equiv \lambda \xi_d$, $\delta_f \equiv \lambda \xi^*$, and $\lambda_t \equiv \lambda u_t$.

References

Andrews, D. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica* 61 (4): 821–56.

Auer, R., K. Degen, and A. M. Fischer. 2010. "Globalization and Inflation in Europe." Working Paper No. 65, Globalization and Monetary Policy Institute, Federal Reserve Bank of Dallas.

Ball, L. 2006. "Has Globalization Changed Inflation?" NBER Working Paper No. 12687.

Ball, L., and R. Moffitt. 2001. "Productivity Growth and the Phillips Curve." NBER Working Paper No. 8421.

Borio, C., and A. Filardo. 2007. "Globalization and Inflation: New Cross-country Evidence on the Global Determinants of Domestic Inflation." BIS Working Paper No. 227.

Buiter, W. H. 2000. "Monetary Misconceptions: New and Old Paradigmata and Other Sad Tales." CEPR Discussion Paper No. 2365.

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

Calza, A. 2008. "Globalization, Domestic Inflation and Global Output Gaps: Evidence from the Euro Area." ECB Working Paper No. 890.

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

Galí, J. 2010. "Inflation Pressures and Monetary Policy in a Global Economy." *International Journal of Central Banking* 6 (1): 93–102.

Galí, J., and M. Gertler. 1999. "Inflation Dynamics: A Structural Econometric Analysis." *Journal of Monetary Economics* 44 (2): 195–222.

Galí, J., and T. Monacelli. 2005. "Monetary Policy and Exchange Rate Volatility in a Small Open Economy." *Review of Economic Studies* 72 (3): 707–34.

Gamber, E., and J. Hung. 2001. "Has the Rise in Globalization Reduced U.S. Inflation in the 1990s?" *Economic Inquiry* 39 (1): 58–73.

Godfrey, L. 1994. "Testing for Serial Correlation by Variable Addition in Dynamic Models Estimated by Instrumental Variables." *Review of Economics and Statistics* 76 (3): 550–59.

Gruben, W., and D. Mcleod. 2004. "The Openness–Inflation Puzzle Revisited." *Applied Economics Letters* 11 (8): 465–68.

Guerrieri, L., C. Gust, and D. López-Salido. 2010. "International Competition and Inflation: A New Keynesian Perspective." *American Economic Journal: Macroeconomics* 2 (4): 247–80.

Hansen, B. E. 1997. "Approximate Asymptotic *P* Values for Structural Change Tests." *Journal of Business and Economic Statistics* 15 (1): 60–80.

Hansen, L. P. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50 (4): 1029–54.

Ihrig, J., S. Kamin, D. Lindner, and J. Marquez. 2010. "Some Simple Tests of the Globalization and Inflation Hypothesis." *International Finance* 13 (3): 343–75.

International Monetary Fund. 2006. "How Has Globalization Changed Inflation?" In *World Economic Outlook: Globalization and Inflation.* Washington, D.C.: International Monetary Fund.

Kamin, S. B., M. Marazzi, and J. W. Schindler. 2004. "Is China 'Exporting Deflation'?" International Finance Discussion Paper No. 791, Board of Governors of the Federal Reserve System.

Lopez-Villavicencio, A., and S. Saglio. 2014. "Is Globalization Weakening the Inflation–Output Relationship?" *Reviews of International Economics* 22 (4): 744–58.

Martínez-García, E. 2015. "Global Slack as a Determinant of US Inflation." Working Paper No. 123, Globalization and Monetary Policy Institute, Federal Reserve Bank of Dallas.

Milani, F. 2010. "Global Slack and Domestic Inflation Rates: A Structural Investigation for G-7 Countries." *Journal of Macroeconomics* 32 (4): 968–81.

———. 2012. "Has Globalization Transformed U.S. Macroeconomic Dynamics?" *Macroeconomic Dynamics* 16 (2): 204–29.

Mishkin, F. 2006. *The Next Great Globalization.* Princeton, NJ: Princeton University Press.

———. 2007. "Inflation Dynamics." *International Finance* 10 (3): 317–34.

———. 2009. "Globalization, Macroeconomic Performance, and Monetary Policy." *Journal of Money, Credit and Banking* 41 (s1): 187–96.

Obstfeld, M., and K. Rogoff. 1995. "Exchange Rate Dynamics Redux." *Journal of Political Economy* 103 (3): 624–60.

———. 2000. "New Directions for Stochastic Open Economy Models." *Journal of International Economics* 50 (1): 117–53.

O'Rourke, K., and G. Williamson. 1999. *Globalization and History: The Evolution of a Nineteenth-Century Atlantic Economy.* Cambridge, MA: MIT Press.

Pagan, A. 1984. "Econometric Issues in the Analysis of Regressions with Generated Regressors." *International Economic Review* 25 (1): 221–47.

Pain, N., I. Koske, and M. Sollie. 2006. "Globalisation and Inflation in the OECD Economies." Working Paper No. 524, OECD Economics Department.

Pehnelt, G. 2007. "Globalization and Inflation in OECD Countries." Working Paper No. 04/2007, ECIPE.

Roberts, J. 2006. "Monetary Policy and Inflation Dynamics." *International Journal of Central Banking* 2 (3): 193–230.

Romer, D. 1993. "Openness and Inflation: Theory and Evidence" *Quarterly Journal of Economics* 108 (4): 869–903.

Smets, F., and R. Wouters. 2002. "Openness, Imperfect Exchange Rate Pass-through and Monetary Policy." *Journal of Monetary Economics* 49 (5): 947–81.

Stock, J. H., and M. W. Watson. 2007. "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking* 39 (s1): 3–33.

Stock, J. H., and M. Yogo. 2003. "Testing for Weak Instruments in Linear IV Regression." NBER Technical Working Paper No. 284.

Temple, J. 2002. "Openness, Inflation, and the Phillips Curve: A Puzzle." *Journal of Money, Credit and Banking* 34 (2): 450–68.

Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, NJ: Princeton University Press.

———. 2009. "Globalization and Monetary Control." In *International Dimensions of Monetary Policy*, ed. J. Galí and M. Gertler, 13–77. Chicago: University of Chicago Press.

Zhang, C., D. Osborn, and D. Kim. 2008. "The New Keynesian Phillips Curve: From Sticky Inflation to Sticky Prices." *Journal of Money, Credit and Banking* 40 (4): 667–99.

Quantitative Easing and Tapering Uncertainty: Evidence from Twitter*

Annette Meinusch and Peter Tillmann
Justus-Liebig-University Gießen, Germany

In this paper we analyze the extent to which people's changing beliefs about the timing of the exit from quantitative easing ("tapering") affect asset prices. To quantify beliefs of market participants, we use data from Twitter, the social media application. Our data set covers the entire Twitter volume on Federal Reserve tapering in 2013. Based on the time series of beliefs about an early or late tapering, we estimate a structural VAR-X model under appropriate sign restrictions on the impulse responses to identify a belief shock. The results show that shocks to tapering beliefs have non-negligible effects on interest rates and exchange rates. We also derive measures of monetary policy uncertainty and disagreement of beliefs, respectively, and estimate their impact. The paper is one of the first to use social media data for analyzing monetary policy and also adds to the rapidly growing literature on macroeconomic uncertainty shocks.

JEL Codes: E32, E44, E52.

1. Introduction

After the 2008 economic crisis, the U.S. Federal Reserve adopted a series of unconventional monetary policy measures in order to

*We thank seminar participants at the Verein für Socialpolitik (Münster 2015), the European Economic Association (Mannheim 2015), the Money, Macro and Finance Group (Cardiff 2015), the 9th International Conference on Computational and Financial Econometrics (London 2015), the Deutsche Bundesbank Big Data Group, the CFCM conference (Nottingham 2016), and the Universities of Duesseldorf and Giessen for important feedback. We are grateful to Michael McMahon and Matthias Neuenkirch for helpful comments. Lucas Hafemann and Florian Viereck provided invaluable research assistance. Author e-mails: annette.meinusch@wirtschaft.uni-giessen.de and peter.tillmann@wirtschaft.uni-giessen.de.

enhance credit conditions and support economic recovery. Large-scale asset purchases known as quantitative easing (QE) led to a tripling of the Federal Reserve's balance sheet.[1] When Federal Reserve Chairman Ben Bernanke, while testifying before the U.S. Congress, first mentioned the possibility of reducing asset purchases on May 22, 2013, markets were wrong-footed. Bernanke's remarks triggered fears of a premature end of asset purchases and an earlier than expected increase in the federal funds rate. Markets coined the term "tapering" to describe the reduction of asset purchases by the Federal Reserve and the eventual end of QE.

Market jitters following the May 22, 2013 testimony led to a sharp increase in long-term interest rates in the United States, a period of high volatility on asset markets, and a dramatic appreciation of the U.S. dollar, particularly against emerging market currencies. Since a large part of these turbulences appeared exaggerated and panic-driven, observers referred to the "taper tantrum."

Federal Reserve Governor Jeremy Stein (2014) reflected on the revision of investors' expectations and the strong market movements in 2013 that gave rise to the "tantrum" notion:

> In early May 2013, long-term Treasury yields were in the neighborhood of 1.60 percent. Two months later, shortly after our June 2013 FOMC meeting, they were around 2.70 percent. Clearly, a significant chunk of the move came in response to comments made during this interval by Chairman Bernanke about the future of our asset purchase program. Perhaps it is not surprising that news about the future course of the asset purchase program would have a strong effect on markets. But here is the striking fact: According to the Survey of Primary Dealers conducted by the New York Fed, there was hardly any change over this period in the expectation of the median respondent as to the ultimate size of the program. Chairman Bernanke's comments . . . did not have any clear directional implications for the total amount of accommodation to be provided via asset purchases.[2]

[1]See D'Amico et al. (2012) and Rogers, Scotti, and Wright (2014) for recent surveys on the effectiveness of unconventional monetary policies.

[2]In April 2013, the pessimistic first quartile of institutions asked by the Federal Reserve Bank of New York's Survey of Primary Dealers showed that markets

In this paper, we provide an empirical analysis of the revision of expectations of market participants and its impact on asset prices that gave rise to the taper tantrum. First, we quantify the response of interest rates, exchange rates, and other asset prices to the belief shocks of market participants specific to tapering beliefs. Second, we measure monetary policy uncertainty as well as derive a measure of disagreement of market participants about future monetary policy and quantify their effects on financial variables. Third, we decompose the dynamics of asset prices in order to isolate the fraction of movements due to changes in tapering beliefs.

The primary difficulty of any study addressing sudden changes in beliefs and their consequences is that individual beliefs about the future course of monetary policy are not observable. Survey evidence is typically only available on a very low frequency, thus making an analysis of daily data impossible. An alternative would be to use beliefs extracted from futures prices or the yield curve. The disadvantage is that these market prices do not allow us to extract measures of disagreement of market participants.

In this paper, we offer a new approach to identify shocks to people's beliefs about monetary policy by using social media. We use data from Twitter.com, the popular social media application for short text messages ("tweets" of no more than 140 characters) since many market participants use their Twitter account to express and disseminate their views on the future stance of monetary policy. To the best of our knowledge, Twitter data has not been used to study monetary policy before.

The advantage of using Twitter data for research purposes is that (i) users not only receive information but can actively share information, (ii) tweets reflect personal views of market participants, (iii) tweets can be used to extract not only a consensus view on policy but also the degree of uncertainty and disagreement about policy, and (iv) Twitter users can respond immediately to news about policy such as Bernanke's testimony and also to other Twitter users'

expected the Federal Reserve to reduce its monthly purchases of assets worth 85 billion dollars at its December meeting. The events in May 2013 triggered a reassessment of expectations. In the July survey, market professionals were expecting purchases of only 65 billion dollars at the September FOMC meeting and only 50 billion dollars after the December meeting.

contributions. Our data set allows us to track the evolution of market beliefs about monetary policy up to the second.

We use the entire Twitter volume containing the words "Fed" and "taper," which amounts to almost 90,000 tweets for the period April to October 2013. From this, we identify tweets that express an explicit view about whether the reduction of bond purchases will occur soon or whether it will occur late. The resulting time series of beliefs of early or late tapering are then put into a vector autoregression (VAR-X) with daily data on interest rates and exchange rates and exogenous variables that control for FOMC meeting days and real economic activity.

With appropriate sign restrictions we are able to identify belief shocks and their dynamic effects. In addition, we use Twitter data to construct two indexes reflecting the uncertainty and disagreement of future Federal Reserve policy and estimate the impact of uncertainty shocks in our VAR model.

The results show that "tapering soon" belief shocks lead to a significant increase in long-term interest rates and a persistent appreciation of the U.S. dollar. A prototypical belief shock raises the share of all tweets considering an early tapering by 10 percentage points, and leads to a 3 basis point increase in long-term yields and a 0.2 percent appreciation of the dollar. These results are in line with the considerations of Krishnamurthy and Vissing-Jorgensen (2013). The data also allow us to study the effects of an increase in uncertainty and disagreement among Twitter users on asset prices. Thus, we can shed light on the points raised by Kashyap (2013), stressing the importance of disagreement about the course of tapering unconventional monetary policy.

Understanding market responses to exiting from QE and other unconventional monetary policies is important. Not only is the Federal Reserve about to gradually exit from unconventional monetary policy, but the Bank of England and, at some point in the future, the Bank of Japan and the European Central Bank are all on the brink of similar exits from unconventional monetary policies. Clearly communicating exit strategies from unconventional monetary policy to financial markets participants and the general public is essential for a smooth and frictionless return to normal. Analyzing data from social media is a useful way of cross-checking whether official communication was received by the markets as intended. In addition, it is important to quantify the impact of market beliefs on interest rates in light of forward guidance used by many central banks.

The remainder of the paper is organized into 7 sections. Section 2 briefly reviews the related literature. Section 3 introduces our data set on Twitter messages, which is used for the empirical analysis in section 4. The results are discussed in section 5. Section 6 is devoted to the analysis of extensions and the robustness of our results. Section 7 includes our concluding remarks and draws some policy implications.

2. The Literature on Federal Reserve Tapering and Uncertainty Shocks

This paper is related to various strands of the literature. There are several papers that focus on the impact tapering announcements have on asset prices. A very useful collection of facts related to the responses to tapering news is provided by Sahay et al. (2014).

Eichengreen and Gupta (2015) present the earliest systematic analysis of Federal Reserve tapering. They attribute the fluctuations in emerging economies in 2013 to Federal Reserve tapering and explain the magnitude of fluctuations in terms of initial macroeconomic conditions. It is shown that better macroeconomic fundamentals did not necessarily shield economies from the tapering fallout.

Aizenman, Binici, and Hutchison (2016) estimate a panel model with daily data for emerging economies and relate the response to tapering news to macroeconomic fundamentals. Similar to Eichengreen and Gupta (2015), they show that fundamentally stronger countries were more sensitive to tapering and argue that this is due to the massive capital inflows these countries received under the Federal Reserve's quantitative easing programs. Their paper uses dummies for FOMC meetings during 2013 as a proxy for tapering news.

Nechio (2014) provides descriptive evidence for the adjustment of emerging economies after Chairman Bernanke's May 22 testimony. She finds that the relative strength of emerging markets' responses reflect internal and external weaknesses specific to each market. Daily data on twenty-one emerging countries is used by Mishra et al. (2014). In contrast to Eichengreen and Gupta (2015) and Aizenman, Binici, and Hutchison (2016), their evidence supports the notion that countries with stronger macroeconomic fundamentals experienced a smaller depreciation of their currencies and smaller increases in borrowing costs. In this study, the market responses are measured in a two-day event window around an FOMC meeting or a publication day of FOMC minutes.

All of these papers proxy market expectations about Federal Reserve tapering by impulse dummies reflecting FOMC meetings and Chairman Bernanke's testimony, or by focusing on relatively narrow event windows. They do not measure market expectations directly. This is exactly where our paper adds to the literature. We extract information from Twitter messages to construct a high-frequency indicator of market beliefs. This indicator also reflects changes in policy perception between FOMC meetings and, in particular, mounting uncertainty before FOMC meetings, which cannot be appropriately proxied by meeting dummies.

Closest to this paper is the work by Dahlhaus and Vasishtha (2014) and Matheson and Stavrev (2014). The latter authors estimate a bivariate VAR model for U.S. stock prices and long-term bond yields. Sign restrictions are used to identify a fundamental-based news shock leading to an increase in both variables and a monetary shock implying an opposite response of stock prices and yields. The authors show that in the taper tantrum episode monetary shocks were important initially, while news shocks became important towards the end of 2013. Our research, however, measures market expectations from social media and avoids restricting the asset price response. Dahlhaus and Vasishtha (2014) identify a "policy normalization shock" using sign restrictions as one that raises federal funds futures but leaves current rates unchanged. They show that this shock has a significant impact on the common component of capital flows to emerging economies.

More generally, our paper adds to the growing body of literature concerned with the macroeconomic consequences of uncertainty shocks. In recent years, researchers develop indicators of uncertainty and analyze the data using VAR models. The first researcher to use this methodology was Bloom (2009). He presents a structural model of macroeconomic uncertainty affecting second moments and estimates a VAR model that replicates the theoretical findings. Baker, Bloom, and Davis (2016) focus on uncertainty about future economic policy. They construct an uncertainty index by referring to newspaper articles about uncertainty and show that this index has predictive power for several macroeconomic variables. On a business level, Bachmann, Elstner, and Sims (2013) use German survey data in a VAR model. They find that a heightened degree of uncertainty for businesses correlates to higher unemployment, lower investment, and higher refinancing costs.

The only paper focusing on monetary developments so far is Istrefi and Piloiu (2013). The authors use the Baker, Bloom, and Davis (2016) index of policy uncertainty for the United States, the United Kingdom, Germany, and the euro area and show that within a structural VAR model uncertainty raises long-term inflation expectations. In contrast to most of these contributions, our measure of policy uncertainty based on Twitter information directly addresses specific uncertainty about the future course of monetary policy.

Lately, there has been a growing interest in the use of social media (Twitter, Google, Facebook) as a data source for economic analyses. Among others, Choi and Varian (2009, 2012) use Google Trends data to forecast near-term values of economic indicators such as initial claims for unemployment. In the context of financial markets Da, Engelberg, and Gao (2011) derive a measure of investor attention based on Google search data. Vlastakis and Markellos (2012) find a link between Google keyword searches and stock trading volume and stock return volatility. Dergiades, Milas, and Panagiotidis (2015) have shown that social media provides significant short-run information for the Greek and Irish government bond yield differential. Acemoglu, Hassan, and Tahoun (2014) predict protests of Egypt's Arab Spring by a Twitter-based measure of general discontent about the government in power. Our study extends this field of research and is the first to analyze monetary policy based on Twitter messages.[3]

3. Tapering Beliefs on Twitter

We extract market participants' beliefs and their uncertainty about the future course of monetary policy from Twitter messages. Twitter

[3]Tillmann (2015) studies the transmission of tapering-related belief shocks to emerging market economies.

is becoming more and more popular among financial professionals. It allows them to comment on policy and market events and to distribute their view to either their followers or an even wider audience in real time.

For the purpose of this study, we obtained the entire Twitter traffic between April 15 and October 30, 2013 containing the words "Fed" and "taper" from Gnip.com, a provider of social media data. Since the debate was focused around the "taper" buzzword, we are confident that we did not miss important tweets when using these keywords. The data set includes nearly 90,000 tweets from about 27,000 users located in 135 countries and the exact time they were sent. This is a unique data set to study market views during the tapering tantrum episode. Panel A in figure 1 plots the daily evolution of Twitter traffic over time.[4]

It can be seen that the number of tweets increases around Bernanke's testimony and around each FOMC meeting. The use of Twitter peaks at the September 17/18 FOMC meeting, when the Federal Reserve finally decided to continue its QE policy and not begin tapering. The sample period covers the entire taper tantrum episode and is sufficiently long to perform a VAR analysis. Further, the data set comprises each tweet's text message of at most 140 characters as well as the name, the location, and the number of followers of the Twitter user.[5]

3.1 Beliefs about the Timing of Tapering

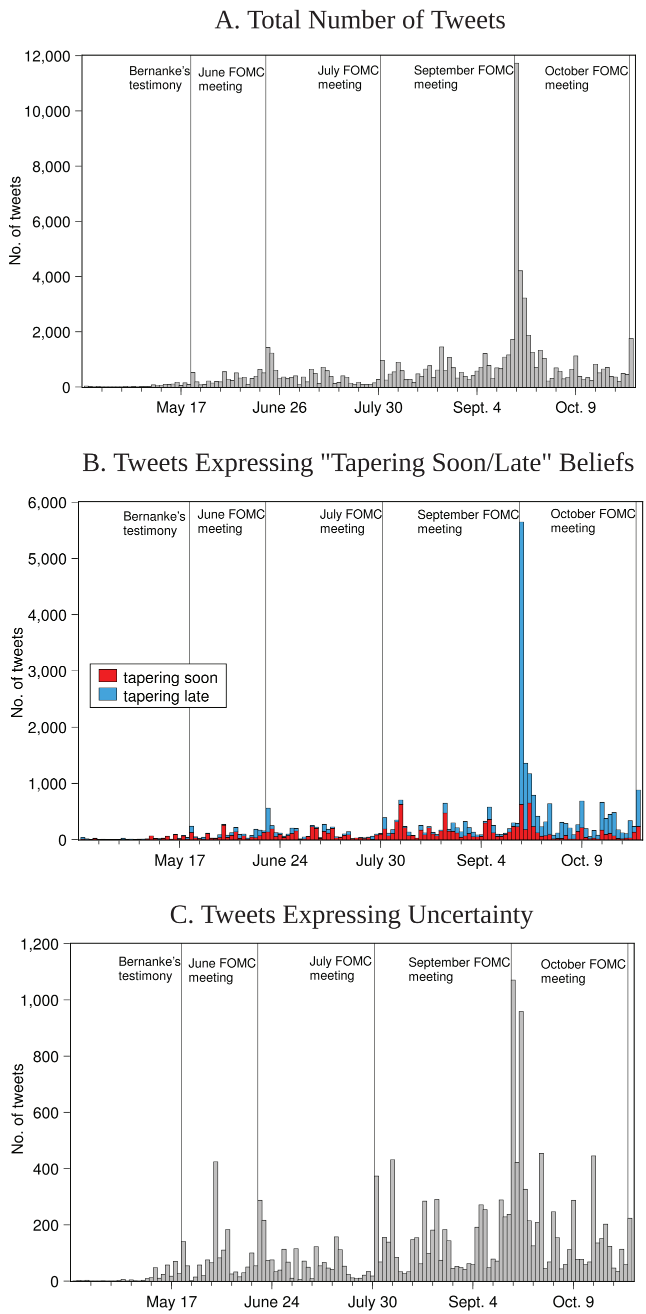
The tweets are separated into those expressing the belief of an early tapering, probably in the summer of 2013 or at the September 2013

---

[4]Retweets, which account for around 26 percent of all tweets, are included in this figure. For the purpose of this paper we interpret retweeted messages as an endorsement of the initial message's relevance and include it in our measure of beliefs. In the robustness section we present results from a model that excludes retweets.

[5]The online appendix (available at https://petertillmann.wordpress.com/research/) contains a plot of the (log) number of users on the ordinate versus the ranked (log) number of tweets on the abscissa. The resulting graph resembles a Zipf-like distribution, indicating that a small number of active users frequently share their opinions about the Federal Reserve's future policy stance and that a large number of users generate tweets about Federal Reserve tapering rather infrequently.

Figure 1. Data Series



Notes: Sample period: April 15–October 30, 2013. The vertical lines indicate FOMC meetings and Chairman Bernanke's May 22 testimony.

FOMC meeting, and those expressing the belief of tapering occurring later. A two-step procedure is used to interpret the content of tweets and allocate the tweets to $Tweets_t^{soon}$ and $Tweets_t^{late}$. In a first step, tweets are filtered according to a list of predefined keywords. In a second step, all remaining tweets are, if appropriate, manually allocated to one of the two categories. The appendix goes into more detail about this procedure. As a result, we are left with roughly 32,000 tweets that contain explicit views about future policy. Tweets that could not be assigned to one of those two categories mostly comment on market movements, point to the upcoming FOMC meeting, or formulate unspecific policy views (i.e., "to taper or not to taper"). Finally, the tweets are aggregated into daily series of beliefs.

As an example, consider the following tweet written on May 20, 2013:

> "Job market gains could lead Fed to taper QE3 early,"

which reflects the view of an early tapering and is allocated to $Tweets_t^{soon}$. Likewise, consider this tweet written on July 31, 2013:

> "The case for a September Fed taper just got a whole lot stronger."

This tweet is also counted as reflecting the view that the Federal Reserve will taper early. The following tweets, in contrast, suggest that Twitter users believe in a later tapering decision. On May 21, 2013, a tweet stated the following:

> "Fed's Bullard says doesn't see a good case for taper unless inflation rises ..."

and on September 18, 2013 the following was retweeted:

> "RT @DailyFXTeam: Economist Nouriel Roubini tweets that based on weak macro data, the Fed shouldn't taper today."

Panel B in figure 1 depicts the identified belief series. We clearly see sizable fluctuations in beliefs and the increased volatility before and after FOMC meetings. Our data contain a total of 15,422 tweets that refer to early tapering and 16,997 tweets that are associated with tapering late. Furthermore, the majority of tweets initially expressed the belief of an early tapering, which then changes in September 2013 in favor of a late tapering. Interestingly, both series, $Tweets_t^{soon}$ and $Tweets_t^{late}$, peak around the September 18 meeting of the FOMC, with a total of 652 “soon” tweets on September 20 and 3,472 “late” tweets on September 18. Further, they are positively correlated with a correlation coefficient of 0.5.

In the regressions below, we include our belief proxies, $Tweets_t^i$, as a fraction of the total amount of timing-related tweets on a particular day:

$$Beliefs_t^i = 100 \times \frac{Tweets_t^i}{Tweets_t^{soon} + Tweets_t^{late}} \ ,$$

with $i = \{soon, late\}$.

3.2 Uncertainty and Disagreement about Tapering

Beliefs of an early or a late tapering could fluctuate within the same day, indicating that there is substantial heterogeneity in market participants' beliefs about future policy. This possibility encourages us to construct two additional indicators that reflect the uncertainty and the disagreement between market commentators about future policy. The uncertainty measure, $Tweets_t^{uncertainty}$, is constructed by counting specific words reflecting uncertainty as in Loughran and McDonald (2011). These authors construct a comprehensive word list to describe uncertainty in text data, which is calibrated to financial applications. Details about the construction are also given in the appendix. Panel C in figure 1 plots the uncertainty indicator. Like the other belief series, uncertainty also seems to be sensitive to official Federal Reserve communication. In particular, uncertainty increases following each FOMC meeting in the sample period.

The indicator $Beliefs_t^{uncertainty}$ is expressed as a ratio of the number of uncertainty-related tweets, $Tweets_t^{uncertainty}$, and the total amount of all tweets on a particular day, i.e.,

$$Beliefs_t^{uncertainty} = 100 \times \frac{Tweets_t^{uncertainty}}{TotalTweets_t} \ .$$

A fourth indicator measures market participants' diverging views about the short-term path of monetary policy. This measure of disagreement, $Beliefs_t^{disagreement}$, is based on soon and late belief series and is defined as

$$Beliefs_t^{disagreement} = 1 - \sqrt{\left(\frac{Tweets_t^{soon}}{TotalTweets_t} - \frac{Tweets_t^{late}}{TotalTweets_t}\right)^2}.$$

It reaches its maximum value of 1 for cases in which the fraction of beliefs corresponding to early tapering is equal to the fraction of beliefs referring to later tapering. If one opinion concerning future monetary policy dominates the other, both fractions diverge and the disagreement index declines. In the following we will use fluctuations in tapering beliefs in a vector autoregressive model to identify unexpected shocks to tapering expectations, policy uncertainty, and investor disagreement.

4. The Model

We use a set of vector autoregressive models with an exogenous variable (VAR-X) to analyze the consequences of shocks to people's beliefs. A combination of sign restrictions is used to identify structural belief shocks.

4.1 The VAR-X Model

Our structural VAR-X model is assumed to have the standard form

$$B(L)Y_t = C + D^{FOMC} + \Theta X_t + \varepsilon_t, \quad \text{with } E[\varepsilon_t \varepsilon_t'] = \Sigma_\varepsilon,$$

where $B(L) \equiv B_0 - B_1 L - B_2 L^2 - ... - B_p L^p$ is a p^{th}-order matrix polynomial in the lag operator L. Further, Y_t is a k-dimensional time series of endogenous variables, X_t is an exogenous variable, and ε_t represents a serially uncorrelated prediction error with Σ_ε as its variance-covariance matrix. The variance-covariance matrix of the structural innovation is normalized to $E(\varepsilon_t \varepsilon_t') \equiv \Sigma_\varepsilon = I_k$.

Since figure 1 shows that our belief series peak on days of FOMC meetings, one could argue that the information content of Twitter beliefs stems from the fact that they simply reflect the official

Federal Reserve communication. To control for this information, we include five dummy variables, D^{FOMC}, for the FOMC meetings in our sample period. For our application it is also important to control for macroeconomic data releases that would affect monetary policy expectations and, as a consequence, asset prices. We include the Aruoba-Diebold-Scotti (2009) daily business conditions index (hereafter the ADS index) as an exogenous variable in order to accomplish this.

A reduced-form representation for this system of equations is

$$A(L)Y_t = \Omega + \Psi X_t + u_t, \quad \text{with } E[u_t u_t'] = \Sigma_u,$$

where $\Omega = B^{-1}(C + D^{FOMC})$ and $\Psi = B^{-1}\Theta$, $A(L) \equiv I - A_1 L - A_2 L^2 - ... - A_p L^p$ reflects the matrix polynomial, and u_t constitutes a white-noise process with variance-covariance matrix Σ_u. Further, the structural shocks are linked to the reduced-form shocks by $u_t = B^{-1}\varepsilon_t$.

The model-specific vector of endogenous variables is

$$Y_t = \left(Beliefs_t^i, \;\; TotalTweets_t, \;\; Rate_t, \;\; FX_t\right)'$$

for $i = (soon, \; late)$ and

$$Y_t = \left(Beliefs_t^j, \;\; Rate_t, \;\; VIX_t, \;\; FX_t\right)'$$

for $j = (uncertainty, \; disagreement)$, where $Rate_t$ is the ten-year constant maturity yield, $TotalTweets_t$ is the log number of daily tweets, and FX_t is the log USD–EUR exchange rate. The VIX index of implied stock market volatility, which is denoted by VIX_t, is needed for identification as discussed below.[6] In order to account for the overall Twitter activity on a particular day, which is a measure of the attention monetary policy receives, we include the total number of tapering tweets. We fit the VAR model to the data by including ten lags of the endogenous variables. All weekends and holidays for which no financial data is available are excluded. The sample period consists of 138 daily observations and covers April 15, 2013 to October 30, 2013 and hence is sufficiently long for reliably estimating a VAR.

[6] All data are taken from the FRED (Federal Reserve Economic Data) database of the Federal Reserve Bank of St. Louis.

4.2 Identification

The identification of belief shocks is crucial for this analysis. As the contemporaneous interaction among all variables at a daily frequency prevents us from using a triangular identification scheme, sign restrictions (Uhlig 2005) provide a useful alternative to identify a structural shock in this VAR analysis. In a sign-restriction approach, identification is achieved by imposing ex post restrictions on the signs of the response of the endogenous variables to a structural shock, e.g., our belief shock. We believe that using sign restrictions creates a VAR best suited to analyze the mutual interaction between market beliefs about policy, asset prices, and volatility indicators even though most of the literature on uncertainty shocks relies on triangular identification schemes instead (such as Bloom 2009 and Baker, Bloom, and Davis 2016, reviewed in section 2).

In order to identify economically meaningful structural shocks, ε_t, we need to find a matrix B_0^{-1} such that the structural innovations are linked to the reduced-form shocks by $u_t = B_0^{-1}\varepsilon_t$, and $\Sigma_u = B_0^{-1}\Sigma_\varepsilon B_0^{-1'} = B_0^{-1}B_0^{-1'}$ with $\Sigma_\varepsilon = I_k$ holds. We proceed in the following way: We estimate our model by OLS, which provides us the reduced-form coefficients $A(L)$ and the covariance matrix Σ_u. Since it is $P_c^{-1} = chol(\Sigma_u)$ so that $\Sigma_u = P_c^{-1}P_c^{-1'}$ and $\Sigma_u = P_c^{-1}\tilde{S}\tilde{S}'P_c^{-1} = B_0^{-1}B_0^{-1'}$ with $B_0^{-1} = P_c^{-1}\tilde{S}$, we randomly draw a matrix $\tilde{S}$ from a space of orthonormal matrices.

Further, we calculate impulse response functions for the restricted periods as $D(L) = A(L)^{-1}B_0^{-1}$ and check whether they satisfy the postulated sign restrictions. We discard those response functions that fail to meet the restrictions while a new orthonormal matrix and new impulse responses are drawn. This procedure is continued until 500 accepted impulse response functions are stored, for which we then compute impulse response functions for all desired periods.

The impact restrictions we use to identify a belief shock are summarized in table 1. We estimate several VAR-X models, one for each alternative series of beliefs and our measures of uncertainty and disagreement. The belief shocks are identified by imposing positive responses of $Beliefs_t^i$, $Beliefs_t^j$, and responses of interest rates and the VIX index. A shock to “tapering soon” beliefs in model I is identified as one that raises the respective belief series and leads to

Table 1. Sign Restrictions to Identify a Belief Shock

	$\boldsymbol{Beliefs_t^i}$	$\boldsymbol{TotalTweets_t}$	$\boldsymbol{Rate_t}$	$\boldsymbol{FX_t}$
Model I: Soon	+		+	
Model II: Late	+		−	
	$\boldsymbol{Beliefs_t^j}$	$\boldsymbol{Rate_t}$	$\boldsymbol{VIX_t}$	$\boldsymbol{FX_t}$
Model III: Uncertainty	+		+	
Model IV: Disagreement	+		+	

higher bond yields. These restrictions are imposed for three periods. The "tapering late" shock raises "tapering late" beliefs and lowers bond yields. We do not restrict the responses for the exchange rate but expect a belief shock to lead to a depreciation of foreign currencies against the U.S. dollar. Since our measure of Twitter beliefs does not include obvious comments on market movements but only firm views on the timing of tapering, we are confident that we can exclude problems of reverse causality.

Since we do not know how shocks to uncertainty and disagreement affect the long-term interest rate, we abstain from restricting those responses in our models III and IV. We assume, however, that both are associated with an increase in market volatility. Hence, we include and restrict the VIX index in our model which is often interpreted as a proxy for fluctuations in risk aversion. As the belief series for uncertainty and disagreement are calculated by using the total amount of tweets on a particular day, we avoid including $TotalTweets_t$ in these two specifications. Although we do not derive the restrictions from a particular asset pricing model, it seems plausible that any increase in policy uncertainty or disagreement among investors is associated with a higher implied volatility. The restrictions for models III and IV are also imposed for three days.

5. Results

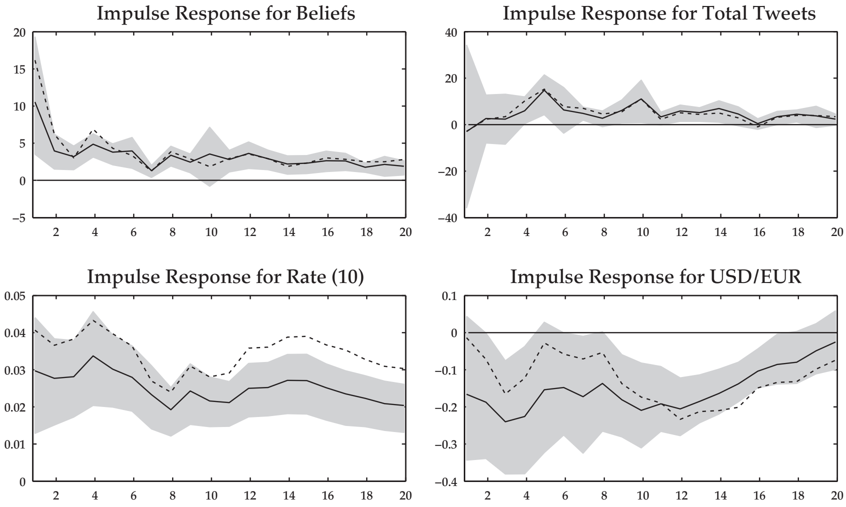
In this section we present the impulse responses following a shock to tapering beliefs, uncertainty, and disagreement about tapering, as identified by the set of sign restrictions. We show the median responses of the variables to a belief shock of one standard deviation for a horizon of twenty days after the shock together with the 16th and 84th percentiles of all accepted impulse responses. Additionally, we show the median-target impulse response (Fry and Pagan 2011), which corresponds to a single structural model that yields an impulse response closest to the median response. We also discuss the forecast error variance decomposition and the historical decomposition of the endogenous variables.

5.1 Shocks to Tapering Beliefs

The responses to a "tapering soon" belief shock are depicted in figure 2. It can be seen that a shift in beliefs towards an early tapering, demonstrated by a 10 percent increase of Twitter users foreseeing an early tapering, leads to a substantial tightening of monetary conditions. Long-term interest rates persistently increase by about 3 basis points. Since some days are characterized by swings in beliefs that are much stronger than the 10 percent depicted here, our model shows that belief shocks could lead to large increases in bond yields. Importantly, our three-day restriction on the direction of the response of bond yields seems to impose a fairly weak constraint on adjustment dynamics. We also find the exchange rate to appear sensitive to tapering beliefs. The dollar appreciates by 0.2 percent following a shift in beliefs. The immediate and persistent response of the dollar to tapering beliefs is in line with the arguments discussed by Krishnamurthy and Vissing-Jorgensen (2013). The total number of tweets, which is a measure of the overall attention of market participants to tapering, adjusts only moderately after a belief shock.

A key feature of our Twitter data set is that we can use the heterogeneity of users' beliefs to distinguish between "tapering soon" and "tapering late" belief shocks. In model II we therefore use the fraction of users expressing a late-tapering belief. Note that by construction, $Tweets_t^{soon}$ and $Tweets_t^{late}$ are not perfectly negatively correlated. Hence, we can estimate the VAR model for late-tapering beliefs in order to compare the strength of the responses to $Tweets_t^{soon}$ and $Tweets_t^{late}$. We obtain inverse results

Figure 2. Impulse Responses to "Tapering Soon" Belief Shock

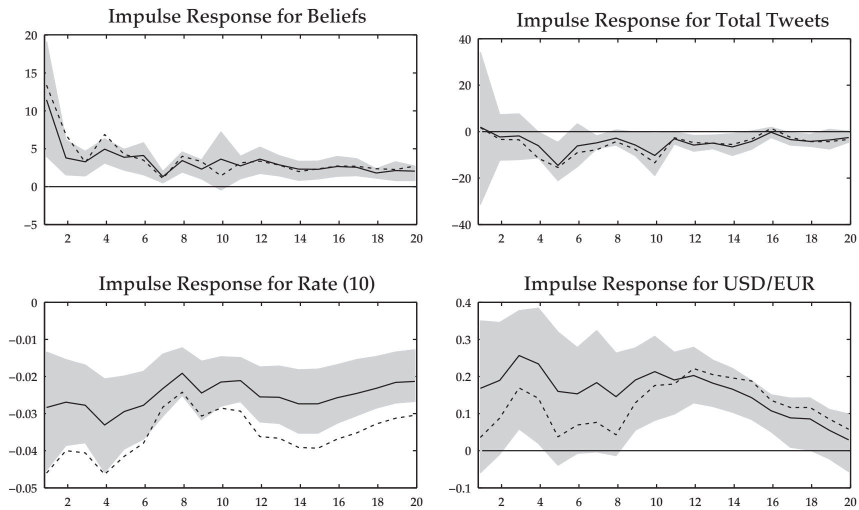


Notes: The solid line is the median impulse response function. The dotted line is the Fry-Pagan (2011) median-target impulse response. The confidence band reflects the 16th and 84th percentiles of all accepted draws.

for responses to a "tapering late" belief shock: the long-term interest rate decreases and the dollar depreciates; see figure 3. Broadly speaking, the responses to "tapering soon" or "tapering late" belief shocks appear symmetric. In both models we appear to underestimate the response of long-term interest rates relative to the Fry-Pagan impulse response.

Figure 4 presents a historical decomposition of interest rates and exchange rates. The contribution of the "tapering soon" belief shock to bond yields, which is depicted by the bars, is particularly pronounced after the June and before the September 2013 FOMC meeting. Likewise, for the exchange rate, the identified shock has large explanatory power before the September meeting. Table 2 decomposes the error variance of the one-, ten-, and twenty-step-ahead forecasts of all variables into the components accounted for by the identified "tapering soon" and "tapering late" belief shocks. For a twenty-day horizon, between 15 percent and 20 percent of the variances of asset prices can be attributed to each shock. Thus, both

Figure 3. Impulse Responses to "Tapering Late" Belief Shock



Notes: The solid line is the median impulse response function. The dotted line is the Fry-Pagan (2011) median-target impulse response. The confidence band reflects the 16th and 84th percentiles of all accepted draws.

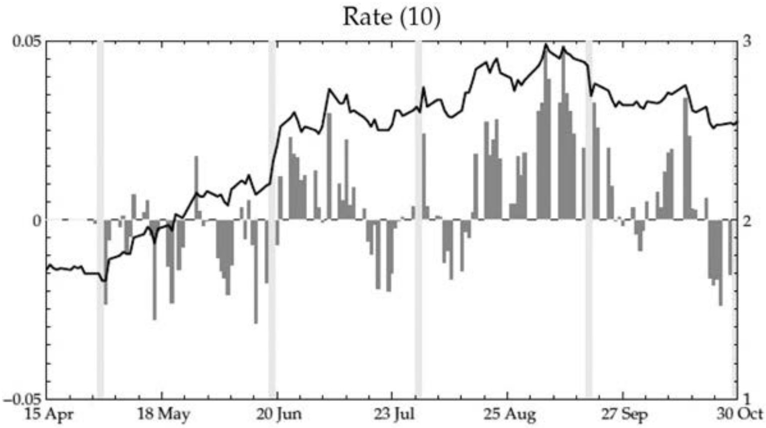
decompositions support the notion of belief shocks as important drivers of bond yields and exchange rates.

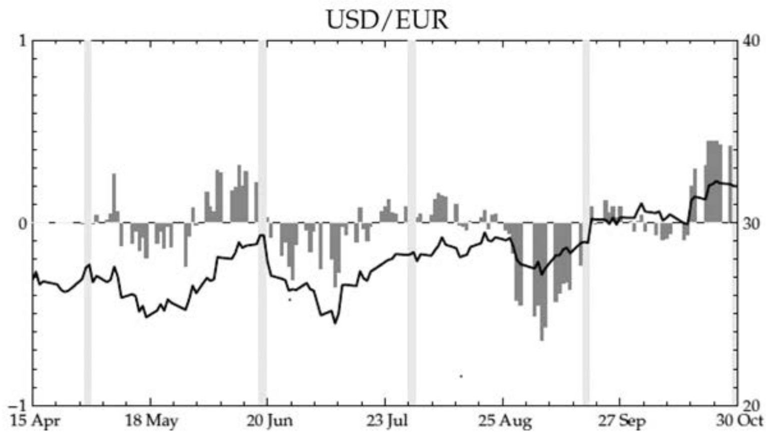
5.2 Shocks to Uncertainty and Disagreement

While the previous subsection studied shifts in the share of Twitter users believing in an early or late tapering, we now analyze the effect of uncertainty and disagreement. Figures 5 and 6 show the results for a shock to the uncertainty and the disagreement indexes, respectively, described in section 3. For these two specifications, no restrictions are imposed on the long-term interest rate and the exchange rate in order to let the data speak freely. However, we restrict the VIX index to respond positively after a belief shock.

We find a persistent response of the VIX index that goes beyond the three-day restriction imposed for identification purposes. Interestingly, an uncertainty shock seems to have no effect on the

Figure 4. Historical Decomposition for "Tapering Soon" Belief Shock





Notes: Contribution of "tapering soon" belief shock (left scale) and observable asset prices (right scale). The shaded areas indicate FOMC meetings.

long-term interest rate.[7] The dollar depreciates by 0.2 percent after an increase in monetary policy uncertainty. A shock to market participants' disagreement about the timing of the tapering decision

[7]Bekaert, Hoerova, and Lo Duca (2013) also use a VAR model and find a negative and persistent effect of uncertainty shocks on the real interest rate.

Table 2. Forecast Error Variance Decomposition

	Impact of Belief Shock (in % of Total Variation)					
	Tapering Soon			Tapering Late		
	At Horizon			At Horizon		
Variable	**1 Day**	**10 Day**	**20 Day**	**1 Day**	**10 Day**	**20 Day**
$Beliefs_t$	23.79	16.77	16.87	18.91	16.25	16.25
$TotalTweets_t$	15.42	14.84	14.81	15.19	14.71	14.65
$Rate_t$	13.02	19.58	19.68	14.36	19.63	19.51
FX_t	14.66	16.98	17.50	13.88	20.10	20.03

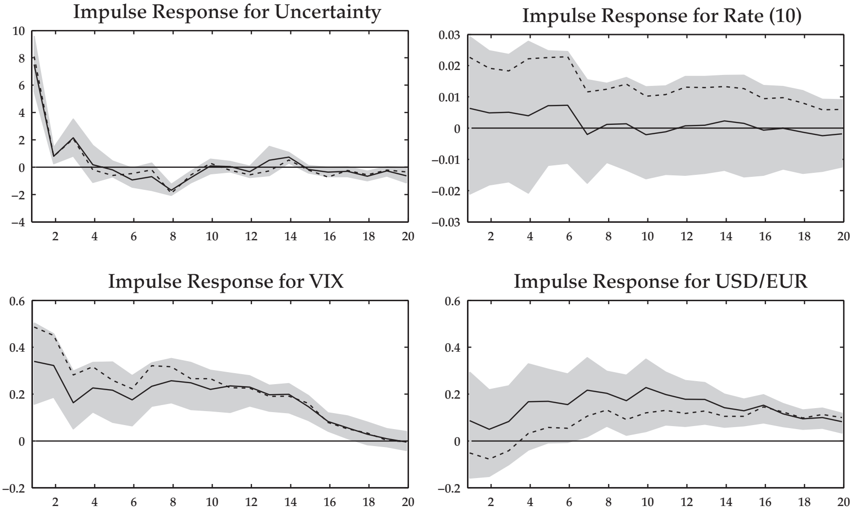
	Impact of Belief Shock (in % of Total Variation)					
	Uncertainty			Disagreement		
	At Horizon			At Horizon		
Variable	**1 Day**	**10 Day**	**20 Day**	**1 Day**	**10 Day**	**20 Day**
$Beliefs_t$	36.54	32.59	32.69	13.81	13.82	14.03
$Rate_t$	12.82	13.40	13.63	17.62	17.36	17.21
VIX_t	12.97	13.52	13.29	17.14	17.13	16.80
FX_t	13.57	14.73	14.69	20.92	19.49	19.32

leads to an increase in interest rates and an appreciation of the dollar. In terms of the signs and the persistence of the responses, these reactions are similar to those after a "tapering soon" shock. Our findings are in line with Kashyap's (2013) view that disagreement is an important explanatory factor for the taper tantrum.

Given that the estimated effects of uncertainty and disagreement diverge with regard to the exchange rate response, we calculate the correlation between the uncertainty and disagreement series. We obtain an unconditional correlation coefficient of 0.03, indicating the absence of a systematic correlation between those two indexes. Thus, we can conclude that shifts in investors' uncertainty or disagreement are separate and quantitatively important factors in the dynamics of interest rates and exchange rates.

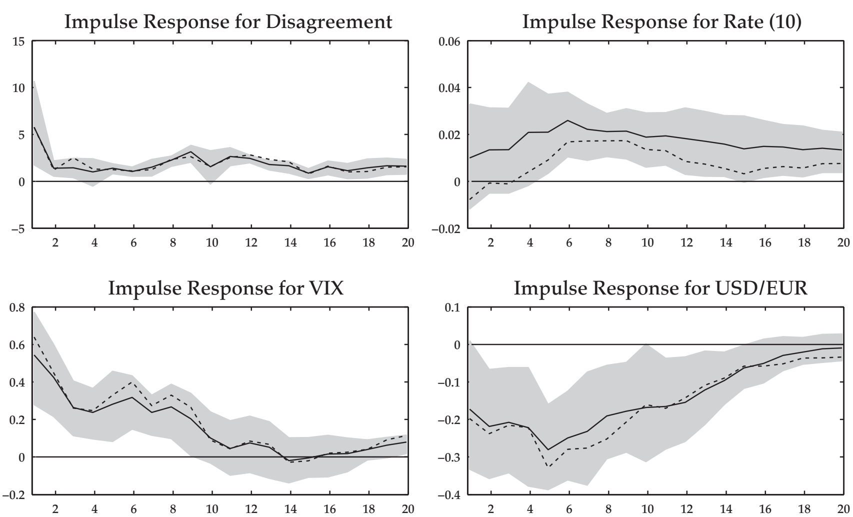
Finally, we report the contribution of uncertainty and disagreement shocks to the forecast error variance of the variables in the

Figure 5. Impulse Responses to Uncertainty Shock



Notes: The solid line is the median impulse response function. The dotted line is the Fry-Pagan (2011) median-target impulse response. The confidence band reflects the 16th and 84th percentiles of all accepted draws.

Figure 6. Impulse Responses to Disagreement Shock



Notes: The solid line is the median impulse response function. The dotted line is the Fry-Pagan (2011) median-target impulse response. The confidence band reflects the 16th and 84th percentiles of all accepted draws.

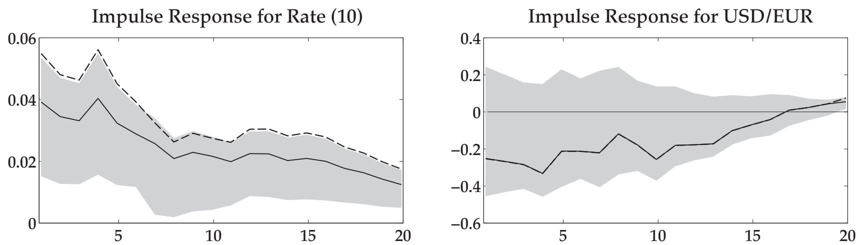
VAR model; see table 2. There is a substantial impact on all included series based on the median estimate of the responses. We find that for all horizons considered, a share between 12 percent and 20 percent of the variation in all variables (for the belief series more than 30 percent) is due to a shock in either variable. This again underlines the quantitative importance of shifts in market beliefs for interest rates and exchange rates.

6. Robustness

In this section, we analyze the robustness of our findings with regard to the role of the information contained in Twitter messages. The first robustness check investigates the role our belief series play in the identification and estimation of shocks. It could be argued that the effects of belief shocks on asset prices entirely stem from the restrictions imposed on bond yields. If the restrictions on the belief series were not binding, the information contained in Twitter would be irrelevant for the result. To shed light on this concern, we estimate a model without Twitter data, that is, we estimate a model that contains the long-term interest rate and the log exchange rate only. The "soon" shock is identified by imposing a positive response of yields. Figure 7 presents the resulting impulse response functions. While both median responses remain unchanged, the exchange rate response is no longer significant. Hence, the information contained in Twitter matters and is indispensable for estimating and identifying powerful shocks.

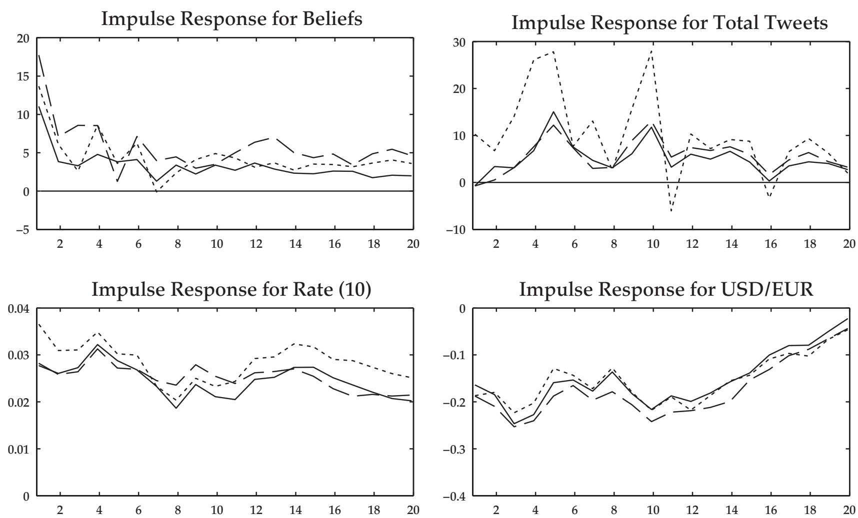
Next, we assess the results of model I by modifying the belief indicators. First, we exclude all retweets from the data set, i.e., forwarded Twitter messages. This is important because it could be argued that retweets do not contain original information and, as a consequence, should not matter for asset prices. We exclude retweets both from $Beliefs_t^i$ and from the total number of tweets. Second, we weight each tweet by the user's number of followers. The relevance of a given tweet should depend on whether the tweet is read by only a few followers or by thousands of subscribers. Again, this is done for $Beliefs_t^i$ and the total number of tweets. Figure 8 shows the impulse responses for both robustness checks. We also report the baseline results for better comparison.

Figure 7. Impulse Responses to "Tapering Soon" Belief Shock: Model Without Twitter Data



Notes: The solid line is the median impulse response function. The dotted line is the Fry-Pagan (2011) median-target impulse response. The confidence band reflects the 16th and 84th percentiles of all accepted draws.

Figure 8. Impulse Responses to "Tapering Soon" Belief Shock: Alternative Twitter Series



Notes: The solid line is the baseline model discussed before. The model that excludes retweets is represented by the dotted line. The dashed line is the result for a model in which tweets are weighted by the number of followers.

It can be seen that our benchmark results for a "sooner" shock (solid line) found in section 5.1 are still robust even when excluding retweeted messages (dotted line). Both specifications of user beliefs deliver almost the same results. This also implies that beliefs expressed through Twitter messages move markets but that the cascade of retweeted messages is of minor importance for financial markets. The results for Twitter series weighted by the number of followers (dashed line) also yield results that hardly differ from the baseline findings. Hence, we can conclude that our results are robust with respect to the way the Twitter data is included.

In addition to these two modifications, we estimate three other models, whose results are not reported in detail. All findings are available in the separate online appendix. In the first model, we include day-of-the-week dummies to account for the fact that new information that is relevant for asset prices is not equally distributed across all weekdays. For example, monetary policy decisions by the European Central Bank and other central banks are typically made on Wednesday or Thursday. FOMC decisions are already captured in the baseline model by appropriate meeting dummies. Likewise, on Monday markets respond to information received over the weekend. This modification leaves all results unchanged.

The second model assesses the identification restrictions. We use a Cholesky ordering of the variables instead of sign restrictions. Tapering beliefs are ordered first such that asset prices can respond contemporaneously to Twitter information. While the results are less significant, their overall magnitude is unchanged.

In a third variation we replace the ten-year bond yields with federal funds rate futures. These are measured by the historical continuous contract data for the CME CBOT thirty-day federal funds futures obtained from Quandl.com. We find that a “tapering soon” belief shock reduces futures prices. This is consistent with an increase in the expected federal funds rate. The response of the exchange rate remains unchanged. A corresponding historical decomposition shows that “soon” beliefs drive futures prices down before the June meeting and around the September FOMC meeting.

7. Conclusions

This paper provides an empirical analysis of the taper tantrum episode of U.S. monetary policy, in which the adjustment of expectations of a normalization of policy caused global market jitters. The analysis is based on a unique data set consisting of 90,000 Twitter messages on Federal Reserve tapering that we use to build series of investors' beliefs about an early or late tapering. A series of VAR estimates showed that shocks to market beliefs derived from Twitter messages have strong and persistent effects on bond yields and exchange rates. The paper is the first study on monetary policy using social media data.

The implications of the findings are threefold. First, our results show that beliefs about exiting QE have contractionary effects on asset prices. This is additional evidence that announcing QE had the intended expansionary effects in the first place.

Second, we show that market sentiment reflected in individual text messages matters for asset prices. Many papers use market prices such as federal funds futures or the yield curve to model expectations of future policy. However, market prices do not allow the researcher to extract information on the uncertainty of the policy outlook or the disagreement among market participants. Twitter data, which we used to show that beliefs of an early or a late tapering could change on the same day, allow such an analysis. Given the ubiquity of social media data and the ability to deal with a large data volume, the usage of this kind of data is an interesting field for future studies in monetary policy.

Third, the study sheds light on the importance of explicitly communicating an exit from unconventional monetary policy measures and offers some quantitative evidence to policymakers. Since many central banks such as the European Central Bank or the Bank of Japan are still heavily engaged in asset purchases and other unconventional policy measures, the challenges of preparing markets for the exit from those policies are yet to come. In this sense the taper tantrum episode of U.S. policy provides valuable lessons that may allow other central banks to avoid exceptional market volatility.

Moreover, our study stands out from many others that analyze the influence of social media on financial markets, because of the uniqueness of our data set. To our knowledge, most of the existing literature relies on, at best, the overall volume only. In comparison with that, our data set comprises each tweet's content, the exact timing the tweet was sent, and the name and location of the Twitter user. Hence, we are able to exploit the data in several dimensions. Given the speed at which news and information spread, it would be interesting to analyze the high-frequency impact of tapering beliefs during and around the FOMC meeting days. This is one task for further research.

Appendix. Construction of Beliefs

Here we describe our procedure of constructing our series of market beliefs, $Tweets^{soon}$ and $Tweets^{late}$, from our set of 87,621 Twitter messages that had been prefiltered out of the entire Twitter traffic by the words “taper” and “Fed.”

We prepare our data set by discarding a small number of tweets written in a language other than English. Then we take into account the fact that tweet data is given in UTC time while all other series, especially asset prices, are based on New York time. Hence, for an adequate estimation of our model, harmonization of the timing is required. Since UTC time is four hours ahead of New York time, we subtract four hours from UTC time to standardize it to New York time. As a consequence, tweets that were posted between 12:00 a.m. and 3:39 a.m. are now assigned to the previous day.

Further, we use a two-step approach to separate beliefs of early tapering from those of late tapering. In a first step, we employ dictionary methods that allow to filter tweets according to a list of predefined keywords. Table 3 and table 4 show the selected keywords for the categories “late” and “soon,” respectively.

It can be seen that both categories are separated into a list of keywords before and after September 18, 2013. This differentiation is necessary because some keywords imply tapering beliefs that depend on the date the corresponding tweet was sent, i.e., a tweet that includes the keyword “December” posted in May corresponds to expectations of a late tapering, while another tweet also referring to “December” but posted in October indicates an early tapering. Keywords that have this property are written in italics. We choose September 18, 2013 as our critical date because of the significant shift in tapering expectations that occurred after the September FOMC meeting, shown by figure 1.

For cases in which the tweets contain negations or keywords from both categories, our dictionary method is not able to allocate tweets to one of the two specified categories. Nevertheless, those tweets are identified by the algorithm that allows us, in a second step, to check and assign them manually.

Table 3. Predefined Keywords for Category "Late"

Late (until September 18, 2013)		Late (from September 19, 2013)	
3rd		*1st*	
2014		2014	
backed away		backed away	
bluff	incl. bluffing	bluff	incl. bluffing
dampen		dampen	
debt ceiling		debt ceiling	
deferred		deferred	
delay	incl. delayed	delay	incl. delayed
December			
dove	incl. dovish	dove	incl. dovish
Dudley		Dudley	
ease fears		ease fears	
end of this year		end of this year	
February		February	
		first	
increase		increase	
		January	
later in 2013		later in 2013	
less	incl. less likely	less	incl. less likely
November			
October			
shutdown		shutdown	
six months		six months	
third			
too soon		too soon	
until		until	
weak	incl. weakness	weak	incl. weakness
will take		will take	

Concerning keywords that are used to identify a series reflecting uncertainty, the reader is referred to Loughran and McDonald (2011). Basically, this procedure is similar to the procedure that is outlined above, except there is no need to create categories before and after September 18 for the specified tweets. For the entire sample it is sufficient to utilize a constant list of keywords.

Table 4. Predefined Keywords for Category "Soon"

Sooner (until September 18, 2013)		**Sooner (from September 19, 2013)**	
		4th	
		2013	
August			
begin		begin	
can taper		can taper	
confidence		confidence	
could taper		could taper	
drop		drop	
early		early	
end eas	incl. end easing	end eas	incl. end easing
expects to taper		expects to taper	
exit qe		exit qe	
fall		fall	
faster		faster	
		fourth	
fuel		fuel	
fell		fell	
Fisher		Fisher	
good news		good news	
hawk		hawk	
increasing expectations		increasing expectations	
in next		in next	
July			
June			
Lacker		Lacker	
likely		likely	
low unemployment		low unemployment	
lower unemployment		lower unemployment	
may begin		may begin	
may soon		may soon	
may taper		may taper	
midyear		midyear	
next few		next gew	
next meeting		next meeting	
		November	
		Nov	
now taper		now taper	
ought to taper		ought to taper	
		Octaper	
		October	
Plosser		Plosser	
pressure		pressure	
quicker		quicker	

(continued)

Table 4. (Continued)

Sooner (until September 18, 2013)		**Sooner (from September 19, 2013)**	
ready		ready	
reduce		reduce	
refine	incl. refining	refine	incl. refining
rumor		rumor	
septaper			
September			
set to taper		set to taper	
should taper		should taper	
slow down		slow down	
soon	incl. sooner	soon	incl. sooner
soonish		soonish	
still		still	
summer			
talk ongoing		talk ongoing	
taper hint		taper hint	
taper sooner		taper sooner	
taper talk		taper talk	
this summer			
unemployment drops		unemployment drops	
unemployment falls		unemployment falls	
unemployment fell		unemployment fell	
urge	incl. urged	urge	incl. urged
will taper off		will taper off	
will taper QE		will taper QE	
within months		within months	
would taper		would taper	
		December	

Although the dictionary approach to the content analysis of tweets is not immune to mistakes, we believe that in the aggregate the resulting belief series are representative for the true beliefs of Twitter users. For every wrongly coded "soon taper" belief, there might be a wrongly coded "late taper" belief. Aggregated over the day, these errors will potentially offset.

References

Acemoglu, D., T. A. Hassan, and A. Tahoun. 2014. "The Power of the Street: Evidence from Egypt's Arab Spring." NBER Working Paper No. 20665.

Aizenman, J., M. Binici, and M. H. Hutchison. 2016. "The Transmission of Federal Reserve Tapering News to Emerging Financial Markets." *International Journal of Central Banking* 12 (2, June): 317–56.

Aruoba, S. B., F. X. Diebold, and C. Scotti. 2009. "Real-Time Measurement of Business Conditions." *Journal of Business and Economic Statistics* 27 (4): 417–27.

Bachmann, R., S. Elstner, and E. Sims. 2013. "Uncertainty and Economic Activity: Evidence from Business Survey Data." *American Economic Journal: Macroeconomics* 5 (2): 217–49.

Baker, S. R., N. Bloom, and S. J. Davis. 2016. "Measuring Economic Policy Uncertainty." Unpublished, Stanford University.

Bekaert, G., M. Hoerova, and M. Lo Duca. 2013. "Risk, Uncertainty and Monetary Policy." *Journal of Monetary Economics* 60 (7): 771–98.

Bloom, N. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77 (3): 623–85.

Choi, H., and H. Varian. 2009. "Predicting Initial Claims for Unemployment Benefits." Technical Report, Google.com.

———. 2012. "Predicting the Present with Google Trends." *Economic Record* 88 (s1): 2–9.

Da, Z., J. Engelberg, and P. Gao. 2011. "In Search of Attention." *Journal of Finance* 66 (5): 1461–99.

D'Amico, S., W. English, D. Lopez-Salido, and E. Nelson. 2012. "The Federal Reserve's Large-Scale Asset Purchase Programs: Rationale and Effects." Finance and Economics Discussion Series Paper No. 2012-85, Board of Governors of the Federal Reserve System.

Dahlhaus, T., and G. Vasishtha. 2014. "The Impact of U.S. Monetary Policy Normalization on Capital Flows to Emerging-Market Economies." Working Paper No. 2014-53, Bank of Canada.

Dergiades, T., C. Milas, and T. Panagiotidis. 2015. "Tweets, Google Trends, and Sovereign Spreads in the GIIPS." *Oxford Economic Papers* 67 (2): 406–32.

Eichengreen, B., and P. Gupta. 2015. "Tapering Talk: The Impact of Expectations of Reduced Federal Reserve Security Purchases on Emerging Markets." *Emerging Markets Review* 25: 1–15.

Fry, R., and A. Pagan. 2011. "Sign Restrictions in Structural Vector Autoregressions: A Critical Review." *Journal of Economic Literature* 49 (4): 938–60.

Istrefi, K., and A. Piloiu. 2013. "Economic Policy Uncertainty and Inflation Expectations." Unpublished, Goethe University Frankfurt.

Kashyap, A. 2013. "Commentary: The Ins and Outs of LSAPs." Proceedings of the 2013 Economic Policy Symposium on "Global Dimensions of Unconventional Monetary Policy," 113–24. Federal Reserve Bank of Kansas City.

Krishnamurthy, A., and A. Vissing-Jorgensen. 2013. "The Ins and Outs of LSAPs." In *Global Dimensions of Unconventional Monetary Policy*, 57–111. Proceedings of the 2013 Economic Policy Symposium sponsored by the Federal Reserve Bank of Kansas City, held in Jackson Hole, Wyoming, August 21–23.

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

Matheson, T., and E. Stavrev. 2014. "News and Monetary Shocks at a High Frequency: A Simple Approach." *Economics Letters* 125 (2): 282–86.

Mishra, P., K. Moriyama, P. N'Diaye, and L. Nguyen. 2014. "Impact of Fed Tapering Announcements on Emerging Markets." IMF Working Paper No. 14/109.

Nechio, F. 2014. "Fed Tapering News and Emerging Markets." Economic Letter No. 2014-06, Federal Reserve Bank of San Francisco.

Rogers, J. H., C. Scotti, and J. H. Wright. 2014. "Evaluating Asset-Market Effects of Unconventional Monetary Policy: A Multicountry Review." *Economic Policy* 29 (80): 749–99.

Sahay, R., V. Arora, T. Arvanitis, H. Faruquee, P. N'Diaye, and T. Mancini-Griffoli. 2014. "Emerging Market Volatility: Lessons from the Taper Tantrum." IMF Staff Discussion Note No. 14/09.

Stein, J. C. 2014. "Challenges for Monetary Policy Communication." Speech at the Money Marketeers of New York University, New York City, May 6, 2014.

Tillmann, P. 2015. "Tapering Talk on Twitter and the Transmission to Emerging Economies." Unpublished (October).

Uhlig, H. 2005. "What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure." *Journal of Monetary Economics* 52 (2): 381–419.

Vlastakis, N., and R. N. Markellos. 2012. "Information Demand and Stock Market Volatility." *Journal of Banking and Finance* 36 (6): 1808–21.

Fiscal Consolidation in an Open Economy with Sovereign Premia and without Monetary Policy Independence*

Apostolis Philippopoulos,[a,b] Petros Varthalitis,[c] and Vanghelis Vassilatos[a]
[a]Athens University of Economics and Business
[b]CESifo
[c]Economic and Social Research Institute, Trinity College Dublin

We welfare rank various tax-spending-debt policies in a New Keynesian model of a small open economy featuring sovereign interest rate premia and loss of monetary policy independence. When we compute optimized state-contingent policy rules, our results are as follows: (i) Debt consolidation comes at a short-term pain, but the medium- and long-term gains can be substantial. (ii) In the early phase of pain, the best fiscal policy mix is to cut public consumption spending to address the debt problem and, at the same time, to cut income tax rates to mitigate the recessionary effects of debt consolidation. (iii) In the long run, the best way of using the fiscal space created is to reduce capital taxes.

JEL Codes: E6, F3, H6.

*We are grateful to Stephanie Schmitt-Grohé (co-editor) and three anonymous referees for constructive criticisms and suggestions. We thank Fabrice Collard, Harris Dellas, Dimitris Papageorgiou, and Johannes Pfeifer for many discussions and comments. We have benefited from comments by seminar participants at the University of Bern, University of Zurich, City University in London, the CESifo Area Conference on Macroeconomics in Munich, and the Athens University of Economics and Business. The first author thanks the Bank of Greece for its hospitality and financial support when this project started. The second author is grateful to the Irakleitos Research Program for financing his doctoral studies. Any views and errors are ours. Corresponding author: Apostolis Philippopoulos, Department of Economics, Athens University of Economics and Business, Athens 10434, Greece. Tel: +30-210-8203357. E-mail: aphil@aueb.gr.

1. Introduction

Since the global crisis in 2008, and after years of deficits and rising debt levels, public finances have been at the center of attention in most euro-zone periphery countries. Although several policy proposals are under discussion, a particularly debated one is public debt consolidation.[1] Proponents claim this is for good reason: as a result of high and rising public debt, borrowing costs have increased, causing crowding-out problems and undermining government solvency. Opponents, on the other hand, claim that debt consolidation worsens the economic downturn and leads to a vicious cycle at least in the short term. At the same time, as members of the single currency, these countries cannot use an independent monetary policy.

What is the best use of fiscal policy under these circumstances? Is debt consolidation beneficial? Should the debt ratio be stabilized at its currently historically high level or should it be brought down? If brought down, how quickly? Do the answers to these questions depend on which tax–spending policy instruments are used over time?

This paper welfare ranks various fiscal policies in light of the above. The setup is a rather conventional New Keynesian model of a small open economy, where the interest rate at which the country borrows from the world capital market increases with the public debt-to-GDP ratio.[2] We focus on a monetary policy regime in which the small open economy fixes the exchange rate and loses monetary policy independence; this mimics membership in a currency union. Hence, the key national macroeconomic tool left is fiscal policy.

Then, following a rule-like approach to policy, we assume that fiscal policy is conducted via simple and implementable feedback policy rules. In particular, we assume that public spending and the tax rates on consumption, capital, and labor are all allowed to respond to the inherited public debt-to-GDP ratio, as well as to contemporaneous

[1] We will use the terms debt consolidation, fiscal adjustment, and fiscal austerity interchangeably. For a discussion of the tradeoffs faced by policymakers in the case of fiscal adjustment, see, e.g., the EEAG Report on the European Economy (European Economic Advisory Group 2014).

[2] For empirical support of this assumption, see, e.g., European Commission (2012). For the small open-economy model and various deviations from it, see Schmitt-Grohé and Uribe (2003). Further details and extensions are below.

output, as deviations from policy targets.[3] We experiment with various policy target values depending on whether policymakers aim just to stabilize the economy around its status quo or whether they also want to move the economy to a new reformed steady state. The status quo is naturally defined as the solution consistent with the euro-period data. The new reformed steady state, on the other hand, is defined as the case in which the fiscal authorities adjust their policies as much as needed so as to end up with lower debt and zero sovereign interest rate premia; we also consider the case in which the new reformed steady state is the associated Ramsey steady state. In addition, since we do not want our results to be driven by ad hoc differences in feedback policy coefficients across different policy rules, we focus on optimized ones. In other words, we compute simple and implementable policy rules that also maximize households' welfare. In particular, adopting the methodology of Schmitt-Grohé and Uribe (2004b, 2006, 2007), we compute welfare-maximizing rules by taking a second-order approximation to both the equilbrium conditions and the welfare criterion around the new reformed steady state(s).

The model is solved numerically using common parameter values and fiscal-public finance data from the Italian economy during 2001–13. We choose Italy simply because it exhibits most of the features discussed in the opening paragraph above and, at the same time, it continues to participate in the world capital market without receiving foreign aid like other euro-zone periphery countries. It thus seems a natural choice to quantify our model.

Before presenting our results, it is worth pointing out that there is no such thing as "the" debt consolidation: the implications of debt consolidation depend heavily on which policy instrument bears the cost in the early phase of austerity and on which policy instrument is anticipated to reap the benefit in the late phase, once the

[3]For empirical support of such simple rules, see, e.g., European Commission (2011). There is a rich literature on monetary and fiscal feedback policy rules that includes, e.g., Schmitt-Grohé and Uribe (2006, 2007), Kirsanova et al. (2007), Pappa and Vassilatos (2007), Batini, Levine, and Pearlman (2008), Leith and Wren-Lewis (2008), Kirsanova, Leith, and Wren-Lewis (2009), Leeper, Plante, and Traum (2009), Bi (2010), Bi and Kumhof (2011), Cantore et al. (2012), Kirsanova and Wren-Lewis (2012), Herz and Hohberger (2013), Kliem and Kriwoluzky (2014), and Philippopoulos, Varthalitis, and Vassilatos (2015).

debt burden has been reduced and fiscal space has been created.[4] The costs in the early phase are due to spending cuts and/or tax increases, while the opposite holds once fiscal space has been created. Our results (see below) confirm all this. Hence, the choice of fiscal policy instruments matters for lifetime utility and output. This choice also matters for how quickly public debt should be brought down: the more distorting are the fiscal policy instruments used during the early costly phase, the slower the speed of fiscal adjustment should be. Naturally, there is more choice when we allow for policy mixes (for instance, when the policy instrument(s) used in the early costly phase can be different from those used in the late phase of fiscal space) than when we are restricted to use a single instrument all the time.

Our main results are as follows. First, in most cases, debt consolidation is beneficial only if we are relatively far-sighted. For instance, in our baseline computations, debt consolidation is welfare-improving only after the first ten years. In other words, debt consolidation comes at a short-term loss (this loss is bigger if one uses one fiscal instrument only, instead of a fiscal policy mix). Nevertheless, once the short-term pain is over, the gains from debt consolidation get substantial over time. All this means that the argument for, or against, debt consolidation involves a value judgment. On the other hand, we find that debt consolidation is welfare-improving all the time, even in the short term, when we travel to the Ramsey steady state; but, in that steady state, the (optimal) values of the tax rates are far away from their values in the actual data.

Second, under debt consolidation, a general result is that the fiscal authorities should use all available tax–spending instruments during the early costly phase of fiscal austerity and reduce capital tax rates—which are particularly distorting—during the late phase of fiscal space. Actually, the anticipation of a reduction in capital taxes plays a key role in the recovery from fiscal austerity. During

[4]In other words, the debate about the benefits and costs of each instrument used for debt consolidation is essentially a debate about the size of the multiplier of each instrument (see the discussion in the EEAG Report on the European Economy (European Economic Advisory Group 2014)). See also, e.g., Coenen, Mohr, and Straub (2008), Leeper, Plante, and Traum (2009), and Davig and Leeper (2011) on how the impact of current policy depends on expectations of possible future policy regimes.

the early costly phase, the assignment of instruments to intermediate targets (or economic indicators) should be as follows: cut public consumption spending to address the public debt problem, and, at the same time, reduce income (capital and labor) tax rates in order to mitigate the recessionary effects of austerity. Sometimes, consumption tax rates should be also used, if changes in other taxes are restricted. The bottom line is that the choice of the fiscal policy mix is important (as also argued by Wren-Lewis 2010) and that the short-term cost becomes too big if the fiscal authorities pay attention to debt imbalances only.

Third, when we solve the model in the fictional case in which Italy would have followed an independent monetary policy (meaning that now there is also a feedback Taylor-type rule for the nominal interest rate), the main results do not change. Also, to the extent that the feedback policy coefficients, both in fiscal and monetary policy rules, are selected optimally, the welfare gain from switching to flexible exchange rates appears to be negligible, at least in this class of New Keynesian models.

What is the value added of our paper? Papers on fiscal consolidation in an open economy, which are close to ours, include Coenen, Mohr, and Straub (2008), Forni, Gerali, and Pisani (2010a, 2010b), Almeida et al. (2013), Cogan et al. (2013), Erceg and Lindé (2013), and Roeger and in 't Veld (2013).[5] But these papers a priori set the fiscal instruments through which debt consolidation is implemented or the speed/pace of this adjustment. Our work differs mainly because: (i) Following an optimized feedback rule type of approach to policy, we search for the best mix of fiscal action in an open economy facing sovereign interest rate premia and loss of monetary policy independence. In doing this, we put special emphasis on which instruments should bear the cost of consolidation in the early phase and which instruments should reap the benefits in the later phase. (ii) We study transition results depending on whether the government simply stabilizes the economy from exogenous shocks

[5]Papers on debt consolidation in a closed economy include Cantore et al. (2012), Bi, Leeper, and Leith (2013), Pappa, Sajedi, and Vella (2015), and Philippopoulos, Varthalitis, and Vassilatos (2015). Econometric studies on the effects of debt consolidation include, e.g., Perotti (1996), Alesina, Favero, and Giavazzi (2012), and Batini, Gallegari, and Giovanni (2012).

or it also leads the economy to a new reformed steady state with lower debt. (iii) We study what would have happened with monetary policy independence, other things being equal.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 presents the data, parameterization, and the status quo solution. Section 4 discusses how we work. The main results are in section 5. A rich sensitivity analysis is in section 6. Section 7 studies the case with independent monetary policy. Section 8 closes the paper. Algebraic details and additional robustness results are in an online appendix.[6]

2. Model

Consider a small open economy where the interest rate premium is debt elastic (see, e.g., Schmitt-Grohé and Uribe 2003). On other dimensions, our setup is the standard New Keynesian model of an open economy with domestic and imported goods featuring imperfect competition and Calvo-type price rigidities (see, e.g., Galí and Monacelli 2005, 2008, and Benigno and Thoenissen 2008).

The economy is composed of N identical households indexed by $i = 1, 2, \ldots, N$; of N firms indexed by $h = 1, 2, \ldots, N$, each one of them producing a differentiated domestically produced tradable good; and of monetary and fiscal authorities. Similarly, there are $f = 1, 2, \ldots, N$ differentiated imported goods produced abroad. Domestic firms are owned by domestic households and any profits are equally divided to these households. Population, N, is constant over time.

2.1 Aggregation and Prices

2.1.1 Consumption Bundles

The quantity of variety h produced by domestic firm h and consumed by domestic household i is denoted as $c^H_{i,t}(h)$. Using a Dixit-Stiglitz aggregator, the composite domestic good consumed by household i,

[6] The appendix is available on the IJCB website (http://www.ijcb.org).

$c_{i,t}^{H}$, consists of h varieties and is given by the function[7]

$$c_{i,t}^{H} = \left[\sum_{h=1}^{N} [c_{i,t}^{H}(h)]^{\frac{\phi-1}{\phi}} \right]^{\frac{\phi}{\phi-1}}, \tag{1}$$

where $\phi > 0$ is the elasticity of substitution across goods.

Similarly, the quantity of imported variety f produced abroad by foreign firm f and consumed by domestic household i is denoted as $c_{i,t}^{F}(f)$. Using a Dixit-Stiglitz aggregator, the composite imported good consumed by household i, $c_{i,t}^{F}$, consists of f varieties and is given by the function

$$c_{i,t}^{F} = \left[\sum_{f=1}^{N} [c_{i,t}^{F}(f)]^{\frac{\phi-1}{\phi}} \right]^{\frac{\phi}{\phi-1}}. \tag{2}$$

In turn, household i's consumption bundle, $c_{i,t}$, is defined as

$$c_{i,t} = \frac{\left(c_{i,t}^{H}\right)^{\nu} \left(c_{i,t}^{F}\right)^{1-\nu}}{\nu^{\nu}(1-\nu)^{1-\nu}}, \tag{3}$$

where ν is the degree of preference for domestic goods (if $\nu > 1/2$, there is a home bias).

2.1.2 Consumption Expenditure, Prices, and Terms of Trade

Household i's total consumption expenditure is

$$P_t c_{i,t} = P_t^{H} c_{i,t}^{H} + P_t^{F} c_{i,t}^{F}, \tag{4}$$

where P_t is the consumer price index (CPI), P_t^{H} is the price index of home tradables, and P_t^{F} is the price index of foreign tradables (expressed in domestic currency).

[7] As in, e.g., Blanchard and Giavazzi (2003), we work with summations rather than with integrals. This does not affect the results.

In turn, i's expenditures on home and foreign goods are, respectively,

$$P_t^H c_{i,t}^H = \sum_{h=1}^{N} P_t^H(h) c_{i,t}^H(h) \tag{5}$$

$$P_t^F c_{i,t}^F = \sum_{f=1}^{N} P_t^F(f) c_{i,t}^F(f), \tag{6}$$

where $P_t^H(h)$ is the price of variety h produced at home and $P_t^F(f)$ is the price of variety f produced abroad, both denominated in domestic currency.

We assume that the law of one price holds, meaning that each tradable good sells at the same price at home and abroad. Thus, $P_t^F(f) = S_t P_t^{H*}(f)$, where S_t is the nominal exchange rate (where an increase in S_t implies a depreciation) and $P_t^{H*}(f)$ is the price of variety f produced abroad denominated in foreign currency. A star denotes the counterpart of a variable or a parameter in the rest of the world. Note that the terms of trade are defined as $\frac{P_t^F}{P_t^H}$ $(=\frac{S_t P_t^{H*}}{P_t^H})$, while the real exchange rate is defined as $\frac{S_t P_t^*}{P_t}$.

2.2 Households

Each household i acts competitively to maximize expected discounted lifetime utility:

$$E_0 \sum_{t=0}^{\infty} \beta^t U\left(c_{i,t}, n_{i,t}, g_t\right), \tag{7}$$

where $c_{i,t}$ is i's consumption bundle as defined above, $n_{i,t}$ is i's hours of work, g_t is per capita public spending, $0 < \beta < 1$ is the time preference rate, and E_0 is the rational expectations operator.

The period utility function is assumed to be of the form[8]

$$u_{i,t}\left(c_{i,t}, n_{i,t}, g_t\right) = \frac{c_{i,t}^{1-\sigma}}{1-\sigma} - \chi_n \frac{n_{i,t}^{1+\eta}}{1+\eta} + \chi_g \frac{g_t^{1-\zeta}}{1-\zeta}, \tag{8}$$

[8]See also Galí (2008), Galí and Monacelli (2008), and many others in this literature.

where χ_n, χ_g, σ, η, and ζ are standard preference parameters. That is, $1/\sigma$ is the elasticity of intertemporal subsititution and η is the inverse of Frisch labor elasticity.

The period budget constraint of each household i written in real terms is (notice that, for simplicity, we assume a cashless economy; we report that our results do not depend on this):

$$\begin{aligned}
&(1+\tau_t^c)\left[\frac{P_t^H}{P_t}c_{i,t}^H + \frac{P_t^F}{P_t}c_{i,t}^F\right] + \frac{P_t^H}{P_t}x_{i,t} + b_{i,t} \\
&\quad + \frac{S_tP_t^*}{P_t}f_{i,t}^h + \frac{\phi^h}{2}\left(\frac{S_tP_t^*}{P_t}f_{i,t}^h - \frac{SP^*}{P}f_i^h\right)^2 \\
&= \left(1-\tau_t^k\right)\left[r_t^k\frac{P_t^H}{P_t}k_{i,t-1} + \widetilde{\omega}_{i,t}\right] + (1-\tau_t^n)\,w_tn_{i,t} \\
&\quad + R_{t-1}\frac{P_{t-1}}{P_t}b_{i,t-1} + Q_{t-1}\frac{S_tP_t^*}{P_t}\frac{P_{t-1}^*}{P_t^*}f_{i,t-1}^h - \tau_{i,t}^l,
\end{aligned} \tag{9}$$

where $x_{i,t}$ is i's domestic investment; $b_{i,t}$ is the real value of i's end-of-period domestic government bonds; $f_{i,t}^h$ is the real value of i's end-of-period internationally traded assets denominated in foreign currency (if negative, it denotes foreign private debt); r_t^k denotes the real return to the beginning-of-period domestic capital, $k_{i,t-1}$; $\widetilde{\omega}_{i,t}$ is i's real dividends received by domestic firms; w_t is the real wage rate; $R_{t-1} \geq 1$ denotes the gross nominal return to domestic government bonds between $t-1$ and t; $Q_{t-1} \geq 1$ denotes the gross nominal return to international assets between $t-1$ and t; $\tau_{i,t}^l$ denotes real lump-sum taxes to each household (if negative, it denotes transfers); and $0 \leq \tau_t^c$, τ_t^k, $\tau_t^n \leq 1$ are tax rates on consumption, capital income, and labor income, respectively. Letters without time subscripts denote steady-state values. The parameter $\phi^h \geq 0$ measures adjustment costs related to private foreign assets as a deviation from their steady-state value, f_i^h; these adjustment costs help us to avoid excess volatility and get plausible (in line with the data) short-term dynamics for private foreign assets following a policy reform; further details are in subsection 3.1 below.

The law of motion of physical capital for each household i is

$$k_{i,t} = (1-\delta)k_{i,t-1} + x_{i,t} - \frac{\xi}{2}\left(\frac{k_{i,t}}{k_{i,t-1}} - 1\right)^2 k_{i,t-1}, \tag{10}$$

where $0 < \delta < 1$ is the depreciation rate of capital and $\xi \geq 0$ is a parameter capturing adjustment costs related to physical capital.

Details on the household's problem and the first-order conditions are in appendix 1.

2.3 Firms

Each firm h produces a differentiated good of variety h, enjoying market power on its own good and facing Calvo-type price fixities.

The profit of each firm h in real terms is (see also, e.g., Benigno and Thoenissen 2008)

$$\widetilde{\omega}_t(h) = \frac{P_t^H(h)}{P_t} y_t^H(h) - \frac{P_t^H}{P_t} r_t^k k_{t-1}(h) - w_t n_t(h). \tag{11}$$

All firms use the same technology as represented by the production function

$$y_t^H(h) = A_t [k_{t-1}(h)]^{\alpha} [n_t(h)]^{1-\alpha}, \tag{12}$$

where A_t is an exogenous stochastic total factor productivity (TFP) process whose motion is defined below and $0 < \alpha < 1$.

Profit maximization by firm h is subject to the demand for its product (see appendix 2):

$$y_t^H(h) = C_t^H(h) + X_t(h) + G_t(h) + C_t^{F*}(h) = \left[\frac{P_t^H(h)}{P_t^H}\right]^{-\phi} Y_t^H, \tag{13}$$

so demand for firm h's product, $y_t^H(h)$, comes from domestic households' consumption and investment, $C_t^H(h)$ and $X(h)$ respectively, where $C_t^H(h) = \sum_{i=1}^{N} c_{i,t}^H(h)$ and $X_t(h) = \sum_{i=1}^{N} x_{i,t}(h)$, from the domestic government, denoted as $G_t(h)$, and from foreign households' consumption, $C_t^{F*}(h) = \sum_{i=1}^{N*} c_{i,t}^{F*}(h)$, and where Y_t^H denotes aggregate demand.

In addition, firms are subject to a Calvo-type pricing mechanism. In particular, in each period, each firm h faces an exogenous probability θ of not being able to reset its price. A firm h, which is able to reset its price at time t, chooses its price $P_t^{\#}(h)$ to maximize the sum of discounted expected nominal profits for the next k periods in which it may have to keep its price fixed. This objective is

$$E_t \sum_{k=0}^{\infty} \theta^k \Xi_{t,t+k} \left\{ P_t^{\#}(h)\, y_{t+k}^H(h) - \Psi_{t+k}\left(y_{t+k}^H(h)\right) \right\},$$

where $\Xi_{t,t+k}$ is a discount factor taken as given by the firm (but it equals the household's intertemporal marginal rate of substitution in consumption, in equilibrium), $y_{t+k}^H(h) = \left[\frac{P_t^{\#}(h)}{P_{t+k}^H}\right]^{-\phi} Y_{t+k}^H$ as said above, and $\Psi_t(h)$ is the minimum nominal cost function for producing $y_t^H(h)$ at t so that $\Psi_t'(h)$ is the associated nominal marginal cost.

Details on the firm's problem and the first-order conditions are in appendix 2.

2.4 Government Budget Constraint

The period budget constraint of the government in real and per capita terms is (details are in appendix 3)

$$\begin{aligned} d_t = {} & R_{t-1}\frac{P_{t-1}}{P_t}\lambda_{t-1}d_{t-1} + Q_{t-1}\frac{S_t P_t^*}{P_t}\frac{P_{t-1}^*}{P_t^*}\frac{P_{t-1}}{P_{t-1}^* S_{t-1}}(1-\lambda_{t-1})\,d_{t-1} \\ & + \frac{P_t^H}{P_t}g_t - \tau_t^c\left(\frac{P_t^H}{P_t}c_t^H + \frac{P_t^F}{P_t}c_t^F\right) - \tau_t^k\left(r_t^k\frac{P_t^H}{P_t}k_{t-1} + \widetilde{\omega}_t\right) \\ & - \tau_t^n w_t n_t - \tau_t^l + \frac{\phi^g}{2}[(1-\lambda_t)\,d_t - (1-\lambda)\,d]^2, \end{aligned} \tag{14}$$

where d_t is the real and per capita value of end-of-period total public debt. Thus, total nominal public debt, D_t, can be held by domestic private agents, $\lambda_t D_t$, as well as by foreign private agents, $(1-\lambda_t)\,D_t$, where the fraction $0 \le \lambda_t \le 1$ is exogenously given.[9] The parameter

[9]Public debt differs from foreign debt. The end-of-period public debt, written in total nominal terms, is $D_t = B_t + S_t F_t^g$, where $B_t = \lambda_t D_t = \sum_{i=1}^N B_{i,t}$ is domestic government bonds held by domestic agents and $S_t F_t^g = (1-\lambda_t)\,D_t$ denotes domestic government bonds held by foreign investors. On the other hand, the country's end-of-period net foreign debt, written in total nominal terms, is $S_t(F_t^g - F_t^h) = (1-\lambda_t)\,D_t - S_t F_t^h$, where $F_t^h = \sum_{i=1}^N F_{i,t}^h$ is nominal foreign

$\phi^g \geq 0$ measures adjustment costs related to public foreign debt; these costs are similar to those of the household in equation (9) above (further details are in subsection 2.7 below).

In each period, one of $(\tau_t^c, \tau_t^k, \tau_t^n, g_t, \tau_t^l, \lambda_t, d_t)$ needs to adjust to satisfy the government budget constraint (see subsection 2.7 below).

2.5 Closing the Model: Debt-Elastic Interest Rate Premium

As is well known, to avoid non-stationarity and convergence to a well-defined steady state, we have to depart from the benchmark small open-economy model (see Schmitt-Grohé and Uribe 2003 for alternative ways). Here, we do so by endogenizing the interest rate faced by the domestic country when it borrows from the world capital market, Q_t.[10] In particular, we start by assuming that the country premium between t and $t+1$, namely $Q_t - Q_t^*$, is an increasing function of the end-of-period total nominal public debt as share of nominal GDP, $\frac{D_t}{P_t^H Y_t^H}$, when this share exceeds a certain threshold. In the robustness section below, we will also study the case where the country premium is increasing in the country's net foreign liabilities or debt.[11]

In particular, following, e.g., Schmitt-Grohé and Uribe (2003) and García-Cicco, Pancrazi, and Uribe (2010), we use

$$Q_t = Q_t^* + \psi \left(e^{\left(\frac{D_t}{P_t^H Y_t^H} - \overline{d} \right)} - 1 \right), \tag{15}$$

assets held by domestic agents (if negative, it denotes liabilities). As said, further details are in appendix 3. Note that, by focusing on a single open economy, we do not model the behavior of foreign investors, so we treat $0 \leq \lambda_t \leq 1$ as exogenous.

[10] See also García-Cicco, Pancrazi, and Uribe (2010), Christiano, Trabandt, and Walentin (2011), and many others for endogenous country premia. Note that although endogeneity of the country premium is also used by the literature on sovereign default, the study of the latter is beyond the scope of our paper. As Corsetti et al. (2013) point out, there are two approaches to sovereign default. The first models it as a strategic choice of the government (Eaton and Gersovitz 1981, Arellano 2008, and many others). The second assumes that default occurs when debt exceeds its endogenous limit (Bi 2012 and many others).

[11] As said above, this rather common assumption (namely, that the interest rate at which the country borrows from the rest of the world is increasing in public and/or foreign debt) is supported by a number of empirical studies (see, e.g., European Commission 2012).

where the world interest rate, Q_t^*, is exogenously given, $\overline{d}$ is an exogenous threshold value above which the interest rate on government debt starts rising above Q_t^*, and the parameter ψ measures the elasticity of the interest rate with respect to deviations of total public debt from its threshold value (see subsection 3.1 below for these parameter values).[12]

2.6 Exchange Rate and Fiscal Policy Regimes

To solve the model, we need to specify the exchange rate and the fiscal policy regimes. Concerning exchange rate policy, since the model is applied to Italy over the last decade, we solve it for a case without monetary policy independence. In particular, we assume that the nominal exchange rate, S_t, is exogenously set and, at the same time, the domestic nominal interest rate on domestic government bonds, R_t, becomes an endogenous variable.[13] Concerning fiscal policy, we assume that, along the transition, the residually determined public financing policy instrument is the end-of-period total public debt, D_t (see below for other public financing cases at the steady state).

Before we turn to fiscal policy rules in the next subsection, it is worth clarifying that, along the transition path, nominal rigidities imply that money is not neutral so that monetary policy, and the exchange rate regime in particular, matters to the real economy.

2.7 Fiscal Policy Rules

Without room for monetary policy independence, only fiscal policy can be used for policy action. In this paper, we follow a rule-like approach to policy. We focus on simple rules, meaning that the fiscal authorities react to a small number of easily observable macroeconomic indicators capturing the current state of the economy.

[12]The value of $\overline{d}$ can be thought of as any value of debt above which sustainability concerns start arising. As we discuss below in subsections 3.1 and 6.1, our qualitative results are robust to the exact value of $\overline{d}$ used, to the extent that each time, the value of ψ is recalibrated.

[13]This is similar to the modeling of, e.g., Erceg and Lindé (2012). Recall that in the case of flexible or managed floating exchange rates, S_t and R_t switch positions, in the sense that S_t becomes an endogenous variable, while R_t is used as a policy instrument usually assumed to follow a Taylor-type rule for the nominal interest rate (see section 7 below).

The magnitude of reaction coefficients (namely, how fiscal policy instruments respond to macroeconomic indicators) will be chosen optimally.

Specifically, we allow the main spending–tax policy instruments (namely, government spending as share of output, s_t^g, and the tax rates on consumption, capital income, and labor income, τ_t^c, τ_t^k, and τ_t^n)[14] to react to the ratio of public liabilities to GDP as a deviation from a target value, $(l_{t-1} - l)$, as well as to the contemporaneous output gap, $\left(y_t^H - y^H\right)$, according to the simple linear rules[15]

$$s_t^g - s^g = -\gamma_l^g \left(l_{t-1} - l\right) - \gamma_y^g \left(y_t^H - y^H\right) \tag{16}$$

$$\tau_t^c - \tau^c = \gamma_l^c \left(l_{t-1} - l\right) + \gamma_y^c \left(y_t^H - y^H\right) \tag{17}$$

$$\tau_t^k - \tau^k = \gamma_l^k \left(l_{t-1} - l\right) + \gamma_y^k \left(y_t^H - y^H\right) \tag{18}$$

$$\tau_t^n - \tau^n = \gamma_l^n \left(l_{t-1} - l\right) + \gamma_y^n \left(y_t^H - y^H\right), \tag{19}$$

where, from the government budget constraint in subsection 2.4, l_{t-1} is defined as

$$l_{t-1} = \frac{R_{t-1}\lambda_{t-1}d_{t-1} + Q_{t-1}\frac{S_t}{S_{t-1}}(1-\lambda_{t-1})d_{t-1}}{\frac{P_{t-1}^H}{P_{t-1}}y_{t-1}^H}$$

and where, in the above rules, l_t, y_t^H, and d_t are written in real and per capita terms, variables without time subscripts denote policy target values, and $\gamma_l^q, \gamma_y^q \geq 0$ for $q \equiv (g, c, k, n)$ are feedback policy coefficients on public debt to GDP and output targets, respectively. The rest of the fiscal policy instruments (namely, lump-sum government tranfers as share of output, denoted as s_t^l, and the share of domestic public debt in total public debt, λ_t) are assumed to be constant over time and equal to their average values in the data (see subsection 2.8 below).

[14] We focus on "distorting" policy instruments, because using lump-sum instruments (such as τ_t^l in our model) to bring public debt down would be like a free lunch.

[15] For similar rules, see, e.g., Schmitt-Grohé and Uribe (2007), Bi (2010), and Cantore et al. (2012). As said above, see European Commission (2011) for similar fiscal reaction functions used in practice. On the other hand, see Kliem and Kriwoluzky (2014) for a critical approach.

Notice that, in the above rules, a policy target value (such as s^g, τ^c, τ^k, τ^n, l, y^H) will be the steady-state value of the corresponding variable. This value will depend on whether we are in the status quo economy or in a reformed economy. For example, as further discussed in section 4 below, the debt policy target, l, can be either the average public-debt-to-GDP ratio in the data (this will be the benchmark case without reforms where fiscal policy adjusts so as to keep the public debt ratio at its average value) or it can be set to a value less than in the data (this will be the case of debt consolidation where fiscal policy systematically brings public debt down over time).

Also, keep in mind that below we will also allow for persistence in policy instruments, as well as for the case in which the debt policy target is not fixed but follows an $AR(1)$ rule (see section 6 below).

2.8 Exogenous Variables and Shocks

In this subsection, we define the exogenous variables and, among them, the exogenous stochastic processes that drive extrinsic fluctuations in our model.

We assume that foreign imports or, equivalently, domestic exports, c_t^{F*}, are a function of terms of trade, $TT_t \equiv \frac{P_t^F}{P_t^H}$, where both variables are expressed as deviations from their steady-state values:

$$\frac{c_t^{F*}}{c^{F*}} = \left(\frac{TT_t}{TT}\right)^{\gamma}, \tag{20}$$

where $0 < \gamma < 1$ is a parameter. The idea is that foreign imports rise as the domestic economy becomes more competitive (the value of the exogenous c^{F*} is specified in the calibration subsection 3 below).

Regarding the other rest-of-the-world variables, namely, the exogenous part of the foreign interest rate, Q_t^*, and the gross rate of domestic inflation in the foreign country, $\Pi_t^{H*} = \frac{P_t^{H*}}{P_{t-1}^{H*}}$, we assume that they are constant over time and equal to $Q_t^* = 1.0115$ (which is the data average value—see below) and $\Pi_t^{H*} \equiv 1$ at all t.

Regarding the exogenously set policy instruments, we set the nominal exchange rate S_t at 1 (under fixed exchange rates), while the output share of government tranfers, s_t^l, and the fraction of domestic public debt in total public debt, λ_t, are set at their data average values at all t (see subsection 3.1 below).

Finally, in the main part of the paper, stochasticity comes from shocks to TFP, which follows:

$$\log(A_t) = (1-\rho^a)\log(A) + \rho^a \log(A_{t-1}) + \varepsilon_t^{\alpha}, \tag{21}$$

where $0 < \rho^a < 1$ is a parameter, $\varepsilon_t^a \sim N\left(0, \sigma_a^2\right)$, and, as said above, variables without time subscript denote steady-state values. As we report below in section 6, our main results do not change when we add extra shocks, such as shocks to the world interest rate.

2.9 Decentralized Equilibrium (for Any Feasible Policy)

We now combine all the above equations to present the decentralized equilibrium (DE) which is for any feasible policy. The DE is defined to be a sequence of allocations, prices, and policies such that (i) households maximize utility; (ii) a fraction $(1-\theta)$ of firms maximize profits by choosing an identical price defined as $P_t^{\#}$, while a fraction θ just set prices at their previous period level; (iii) all constraints, including the government budget constraint and the balance of payments, are satisfied; (iv) markets clear; and (v) policymakers follow the feedback rules assumed in subsection 2.7. This DE is given the values of feedback policy coefficients in the policy rules (16)–(19); the exogenous variables, $\{c_t^{F*}, Q_t^*, \Pi_t^{H*}, \epsilon_t,\ s_t^l,\ \lambda_t, A_t\}_{t=0}^{\infty}$, which have been defined in subsection 2.8; and initial conditions for the state variables.

We thus end up with a first-order non-linear dynamic system of thirty-two equations. Algebraic details, the final equilibrium system, and the associated steady state (also used for calibration) are in appendix 4.

2.10 Solution Methodology

We will work as follows. In the next section (section 3), we solve the above model numerically, employing parameter values and data in accordance with the Italian economy over 2001–13. As we shall see, the steady-state solution of this model economy will do well at mimicking the output shares of key macroeconomic variables observed in the Italian data since 2001; hence, we will call it the "status quo steady state." In turn, the next sections will study the transition dynamics, when we depart from this status quo steady state and travel to a new reformed steady state (the latter is defined in section 4.1 below). To compute the transition dynamics, driven by policy reforms (such as debt consolidation policies) and/or exogenous stochastic proccesses (such as TFP shocks), we will take a second-order aproximation of the stochastic problem around the deterministic new reformed steady state.[16] In all policy experiments, the feedback coefficients in the policy rules are computed optimally to maximize the household's welfare.

Further details on the reformed economy and the computational methodology are in section 4 below. The quantitative implementation is in sections 5 and 6. But we first need to solve for the status quo steady state. This is in the next section.

3. Data, Parameterization, and the Status Quo Steady State

This section parameterizes the above model economy using average data from Italy over 2001–13 (the exact end period for each variable may vary depending on data availability) and then presents the resulting steady-state solution. As said, the latter will then serve as a point of departure to study policy reforms.

3.1 Data and Parameter Values

The data are from Eurostat over 2001–13 (in the case of TFP, the data source is Ameco over 1980–2014). The time unit is meant to be a year. In calibrating the model, we assume that the economy is in

[16]A potential criticism might be that an approximate solution might not be reliable because we approximate the equilibrium equations around a new steady state different from the initial one. To address this issue, we have also solved a deterministic version of the model using non-linear, or non-approximate, numerical methods. This means that we repeat the equilibrium equations in each time period until the economy converges to a steady-state position where variable changes are negiligible. We use Dynare for these computations. The key results, namely, the optimal policy mix during the transition, do not change.

the deterministic steady state of the decentralized equilibrium presented above with zero inflation (see appendix 4 for the steady-state system). Recall that, since policy instruments react to deviations of macroeconomic indicators from their steady-state values, feedback policy coefficients do not play any role at the steady state.

The baseline parameter values, as well as the values of the fiscal policy variables, are listed in table 1. As reported in detail in section 6, our results are robust to changes in these parameter values.

The value of the time preference rate, β, follows from setting the gross interest rate at $R = Q = 1.0225$ (the latter is consistent with an interest rate premium of 1.1 percent over the German ten-year bond rate, which is the average value in the data, and with a gross inflation rate equal to one at the steady state).

The value of a implies a labor share, $(1-a)$, equal to 0.62, which is the average value in the Italian data over 2001–13. We employ conventional parameter values, as used by the literature, for the elasticity of intertemporal substitution, $1/\sigma$, the inverse of Frisch labor elasticity, η, and the price elasticity of demand, ϕ, which are as in, e.g., Andrès and Doménech (2006) and Galí (2008). Regarding the preference parameters in the utility function, χ_n is calibrated from the household's labor supply condition, while χ_g is set at 0.1. The price rigidity parameter, θ, is set at 0.5. The value of γ, in equation (20) for foreign imports, is set at 0.9.

In our baseline parameterization, the threshold parameter value of the public-debt-to-GDP ratio above which sovereign interest rate premia emerge, $\overline{d}$, is set at 0.9 (see equation (15)). This value is consistent with evidence provided by, e.g., Reinhart and Rogoff (2010) and Checherita-Westphal and Rother (2012) that, in most advanced economies, the adverse effects of public debt arise when it is around 90–100 percent of GDP. It is also within the range of thresholds for sustainable public debt estimated by the European Commission (2011). In turn, the associated premium parameter, ψ, is calibrated by using equation (15). In other words, assuming a value for the parameter $\overline{d}$, and using data averages over the sample period for the interest rate premium and the ratio of public debt to GDP, the value of ψ follows from equation (15). In our baseline parameterization, the resulting value of ψ is 0.0505, which means that a 1 percentage point increase in the debt-to-GDP ratio leads to an increase in the interest rate premium by 5.05 basis points. Such values are in

Table 1. Baseline Parameter Values and Policy Variables

Parameter	Value	Description
a	0.38	Share of Capital
β	0.9708	Rate of Time Preference
ν	0.5	Home Goods Bias Parameter at Home
δ	0.04	Rate of Capital Depreciation
ϕ	6	Price Elasticity of Demand
η	1	Inverse of Frisch Labor Elasticity
σ	1	Elasticity of Intertemporal Substitution
ν^*	0.5	Home Goods Bias Parameter Abroad
θ	0.5	Price Rigidity Parameter
ψ	0.0505	Interest Rate Premium Parameter
χ_n	3.66	Preference Parameter Related to Work Effort
χ_g	0.1	Preference Parameter Related to Public Spending
$\bar{d}$	0.9	Threshold Parameter of Public Debt as Share of Output
ρ^a	0.9479	Persistence of TFP (1980–2014)
σ_a	0.007636	Standard Deviation of TFP (1980–2014)
γ	0.9	Terms of Trade Elasticity of Foreign Imports
ξ	0.3	Adjustment Cost Parameter on Physical Capital
ϕ^g	0.3	Adjustment Cost Parameter on Foreign Public Debt
ϕ^h	0.3	Adjustment Cost Parameter on Private Foreign Assets/Debt
τ^c	0.1756	Consumption Tax Rate
τ^k	0.3118	Capital Tax Rate
τ^n	0.421	Labor Tax Rate
s^g	0.2222	Government Spending on Goods/Services as Share of GDP
s^l	0.2326	Government Transfers as Share of GDP
λ	0.64	Fraction of Total Public Debt Held by Domestic Agents
$\frac{c^{F*}}{c^F}$	1.01	Exports-to-Imports Ratio

line with empirical findings for OECD countries (see, e.g., Ardagna, Caselli, and Lane 2008). As reported in subsection 6.1 below, our results are robust to changes in $\overline{d}$, to the extent that each time we recalibrate ψ.

The parameters ϕ^h and ϕ^g, measuring adjustment costs associated with changes in private and public foreign assets, respectively (see equations (9) and (14)), are both set to 0.3. As mentioned above, this value gives plausible short-run dynamics for private foreign assets and, in turn, for the country's net foreign debt following a policy reform. Robustness checks for both ϕ^h and ϕ^g are reported in subsection 6.1 below. We also report that a positive value for ϕ^h is needed to give bounded solutions for the welfare when we compute optimized feedback policy rules below. Similarly, the value of ξ measuring capital adjustment costs is set equal to 0.3.

Concerning the exogenous variables, the persistence and standard deviation parameters in the TFP process (21) are set at $\rho^a = 0.9479$ and $\sigma_a = 0.007636$. These values have been estimated by running simple regressions using Ameco data for total factor productivity. As reported below, our results are robust to changes in these parameter values. Regarding the rest-of-the world variables, Π_t^{H*}, Q_t^*, and c_t^{F*}, we set their steady-state values equal to $\Pi^{H*} \equiv 1, Q^* = 1.0115$ (which is the data average) and $\frac{c^{F*}}{c^F} = 1.01$ (which is the ratio of exports to imports in the Italian data).

The steady-state values of fiscal and public finance policy instruments, τ_t^c, τ_t^k, τ_t^n, s_t^g, s_t^l, and λ_t are set at their data averages. In particular, τ^c, τ^k, and τ^n are the effective tax rates on consumption, capital, and labor in the data over 2001–13. Moreover, s^g and s^l, namely, government spending on goods/services and on transfer payments as shares of output, are set at their average values in the data, 0.2222 and 0.2326, respectively. Finally, λ, the fraction of total public debt held by domestic private agents is set at 0.64, which is again its average value in the data during the same period.

3.2 Status Quo Steady State

The steady-state system is in appendix 4. Table 2 presents the numerical solution of this system when we use the parameter values and policy instruments in table 1. In this solution, we treat public debt, d, as the residually determined public financing instrument. In

Table 2. Status Quo Steady-State Solution

Variables	Description	Steady-State Solution	Data
u	Period Utility	0.8217	—
r^k	Real Return to Physical Capital	0.0908	—
w	Real Wage Rate	1.11378	—
n	Hours Worked	0.331281	0.2183
y^H	Output	0.712326	—
TT	Terms of Trade	0.994923	—
$Q - Q^*$	Interest Rate Premium	0.011	0.011
$\frac{TT^{1-\nu}c}{y^H}$	Consumption as Share of GDP	0.6335	0.5961
$\frac{k}{y^H}$	Physical Capital as Share of GDP	3.4872	—
$\frac{TT^{\nu^*}f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.1813	0.1039
$\frac{TT^{1-\nu}d}{y^H}$	Total Public Debt as Share of GDP	1.0965	1.098
$\widetilde{f}$	Country's Net Foreign Debt as Share of GDP	0.2134	0.2109

table 2, we also present some key ratios in the Italian data. Notice that most of the solved ratios, produced endogenously by the model, are meaningful and close to their actual values. For instance, the solution for the country's net foreign debt as share of GDP (denoted as $\widetilde{f}$)[17] is 0.2134; its average value in the data is 0.2109. Also, the solution for total public debt as share of GDP[18] is 1.0965; its average value in the data is 1.098.

In what follows, this steady-state solution will serve as a point of departure to study various policy experiments. Before we move on to reforms, it is worth pointing out that, in the above model, a lower public debt implies a lower sovereign premium, and this

[17]Thus, $\widetilde{f} \equiv \frac{S_t(F_t^g - F_t^h)}{P^H Y^H} = \frac{(1-\lambda_t)D_t - S_t F_t^h}{P_t^H Y_t^H} = \frac{(1-\lambda)TT^{1-\nu}d - TT^{\nu^*}f^h}{y^H}$, where $TT_t \equiv \frac{P_t^F}{P_t^H}$ is the terms of trade. Details are in appendix 4.

[18]This is $\frac{D}{P^H Y^H} \equiv \frac{TT^{1-\nu}d}{y^H}$, where $TT_t \equiv \frac{P_t^F}{P_t^H}$ is the terms of trade. Details are in appendix 4.

leads to higher capital, higher output, and higher welfare; this can rationalize the debt consolidation policies studied in what follows.

4. Description of Policy Experiments

In this section, we define the policy reform studied, how we model debt consolidation, and how we compute optimized feedback policy rules.

4.1 Definition of the Reformed Economy

The main experiment we want to consider in this paper is the case in which the economy departs from the status quo and travels over time to a new reformed state with lower debt. Regarding the status quo, this is provided by the steady-state solution of the model economy calibrated and solved numerically in the previous section (see tables 1 and 2). Regarding the new reformed steady state, this is defined as the case in which the ratio of public debt to GDP is permanently reduced, so that there are no sovereign interest rate premia in the new steady state. In other words, in the new reformed steady state, we set $Q = Q^*$ so that $\frac{TT^{1-\nu}d}{y^H} = \overline{d}$. Specifically, the government reduces the output share of public debt from 1.0965 (which is the status quo solution) to the threshold value of $\overline{d} = 0.9$ corresponding with zero premia. To put it differently, since, in our model, sovereign premia arise whenever the ratio of public debt to output happens to be above the 0.9 threshold, premia are eliminated $(Q = Q^*)$ once such debt reduction has been achieved. As said above, we will conduct robustness exercises for the values of the exogenously set threshold, $\overline{d}$. Also, again as a robustness check, we will consider the case in which the new reformed steady state is the Ramsey steady state, meaning that the policy targets are the Ramsey steady-state values.

In addition, we assume that, in the new reformed steady state, the country's net foreign debt position is zero or, equivalently, that the country ends up with a balanced trade.[19] In other words, in the

[19]For a similar practice (namely, to assume a zero net foreign debt position in the steady state and then check its robustness), see, e.g., Mendoza and Tesar (2005).

new reformed steady state, we set the country's net foreign debt as share of output to zero, $\widetilde{f} = 0$. This means that the country's net foreign debt as share of output is permanently reduced from 0.2134 (which is the status quo solution) to zero. Note that this assumption is not important to our qualitative results (its robustness is checked in subsection 6.1 below) but, if the foreign position of the country is left free in the long run with relatively low interest rates ($Q = Q^*$), private agents have an incentive to overborrow and this leads to unrealistically high values of private foreign debt, $f^h < 0$.

Details of the equilibrium conditions of the reformed economy, the steady state implied by these conditions, as well as numerical solutions of the reformed steady state under unrestricted or restricted $\widetilde{f}$, and under alternative public financing scenarios, are provided in appendix 5.

4.2 Public Debt Consolidation and the Intertemporal Tradeoff

It is widely recognized that debt consolidation implies a tradeoff between short-term fiscal pain and medium-term fiscal gain. In our model, during the early phase of the transition, debt consolidation comes at the cost of higher taxes and/or lower public spending. On the other hand, in the medium and long run, a reduction in the debt burden allows, other things being equal, a cut in tax rates and/or a rise in public spending. Thus, one has to value the early costs of stabilization vis-à-vis the medium- and long-term benefits from the fiscal space created.

This intertemporal tradeoff also implies that the implications of consolidation depend heavily on the public financing policy instruments used, namely, which policy instrument adjusts endogenously to accommodate the exogenous change in fiscal policy (see also, e.g., Leeper, Plante, and Traum 2009, and Davig and Leeper 2011). Specifically, these implications depend both on which policy instrument bears the cost of adjustment in the early period of adjustment and on which policy instrument is anticipated to reap the benefit once consolidation has been achieved. In the policy experiments we consider below, we experiment with fiscal policy mixes, which means that the fiscal authority can use various instruments in the transition and in the steady state. Details are in subsection 5.3 below.

4.3 The Reformed Economy vs. a Reference Regime

To evaluate the implications of debt consolidation as defined in subsection 4.1 above, we need to compare them with a reference regime. Here, we find it natural to use as reference regime the case without debt consolidation, other things being equal. In other words, we will study two scenaria regarding policy action. The first, used as a reference, is the scenario without debt consolidation. Here, the role of policy is only to stabilize the economy against shocks. For instance, say that the economy is hit by an adverse temporary TFP shock, which, as the impulse response functions reveal, produces a contraction in output, a rise in the ratio of public debt to output, and a rise in the sovereign premium. Then, the policy questions are which policy instrument to use and how strong the reaction of policy instruments to deviations from targets should be, in order to maximize household's welfare criterion. Note that, in this case, the values of the policy targets in the feedback rules (16)–(19) are given by the status quo steady-state solution. In other words, in this policy scenario, we depart from, and end up at, the status quo steady-state solution of subsection 3.2 above, so that transition dynamics are driven by exogenous shocks only.

The scenario with debt consolidation is richer. Now the role of policy is twofold: to stabilize the economy against the same shocks as above and, at the same time, to improve resource allocation by gradually reducing the debt-to-GDP ratio over time as defined in subsection 4.1. The policy questions are as above in the reference regime, except that now the policy targets in the feedback rules (16)–(19) are given by the steady-state solution of the new reformed economy. In other words, in this case, we depart from the status quo solution with sovereign premia, but we end up at a new reformed steady state with lower (public and foreign) debt. Thus, now there are two sources of transition dynamics: temporary shocks and the deterministic difference between the initial and the new reformed long run (see also Cantore et al. 2012).

4.4 Computational Methodology

Irrespectively of the policy experiments studied, to make the comparison of different policy regimes meaningful, we compute optimized policy rules, so that our results do not depend on ad hoc differences in feedback policy coefficients across different policy rules. The welfare criterion is the household's expected discounted lifetime utility (see equations (7)–(8) above).

We work as follows: First, we take a second-order approximation of the equilibrium conditions, as well as of household's expected discounted lifetime utility, around the deterministic reformed steady state (see appendix 5 for the latter).[20] This is a function of feedback policy coefficients in the policy rules and initial values for the state variables. Secondly, we select the feedback policy coefficients so as to maximize the conditional mean of household's expected discounted lifetime utility, where conditionality refers to the initial conditions chosen; the latter are given by the status quo steady-state solution (see subsection 3.2 above). Thus, the initial values of the endogenous and exogenous predetermined variables are set equal to their status quo values. If necessary, the ranges of the feedback policy coefficients or, equivalently, the values of the policy instruments themselves will be restricted to give determinate solutions and/or meaningful values for policy instruments (e.g., tax rates less than one and non-negative nominal interest rates). All this is similar to Schmitt-Grohé and Uribe (2004b, 2006, 2007).[21] As said above, we work similarly in the case without debt consolidation (where we depart from, and return to, the same status quo steady state), which serves as a reference regime.

5. Macroeconomic Implications of Fiscal Consolidation

In this section, we present the main results. The emphasis will be on the case of the reformed economy as defined in subsection 4.1

[20]Thus, we take a second-order approximation to both the equilibrium conditions and the welfare criterion. As is known, this is consistent with risk-averse behavior on the part of economic agents and can also help us to avoid possible spurious welfare results that may arise when one takes a second-order approximation to the welfare criterion combined with a first-order approximation to the equilibrium conditions (see, e.g., Gali 2008 and Benigno and Woodford 2012).

[21]Specifically, to compute a second-order accurate approximation of both the conditional welfare and the decentralized equilbrium, as functions of feedback policy coefficients, we use the perturbation method of Schmitt-Grohé and Uribe (2004b). In turn, we use a MATLAB function (such as fminsearch.m) to compute the values of the feedback policy coefficients that maximize this approximation.

Table 3. Steady-State Utility and Output in the Reformed Economy

Residual Fiscal Instrument	Steady-State Utility	Steady-State Output
s^g	0.9125	0.82367
τ^c	0.9227	0.82367
τ^k	0.9311	0.834774
τ^n	0.9290	0.83139
s^l	0.9180	0.817941

above, but, for reasons of comparison, we will also present results for the reference case without debt consolidation. Recall that in the case of debt consolidation, transition dynamics are driven both by a temporary supply shock and by deterministic changes in fiscal policy instruments aiming at debt reduction and elimination of the sovereign premium over time.

5.1 *Steady-State Utility and Output in the Reformed Economy*

The steady state of the reformed economy with debt consolidation is as defined in the previous section, while details are in appendix 5. Thanks to the fiscal space created by debt reduction, public spending can rise or a tax rate can be reduced residually.

Table 3 summarizes steady-state utility and output under alternative public financing scenarios in this reformed economy (for the full numerical solutions, see appendix 5). In the first row of table 3, the assumption is that it is public spending that takes advantage of debt reduction, in the sense that, once the debt burden has been reduced, public spending can increase relative to its value in the status quo solution. In the next three rows, the fiscal space is used to finance cuts in one of the three tax rates. The last row reports, for comparison, the case in which it is lump-sum tranfers, s^l, that rise. The best outome (both in terms of utility and output) is achieved when the fiscal space is used to finance a cut in capital tax rates. This is as expected and is consistent with the Chamley-Judd normative result. Therefore, in what follows, we will use the capital tax rate, τ^k, as the residually determined fiscal policy instrument in the steady state of the reformed economy with debt consolidation.

5.2 Determinacy and Bounds on Policy Coefficients

As is known, local determinacy, or implementability, depends crucially on the values of feedback policy coefficients in the rules (16)–(19) above. This is also the case in our model. Our experiments show that economic policy can ensure determinacy when at least one of the fiscal policy instruments (s_t^g, τ_t^c, τ_t^k, τ_t^n) reacts to public liabilities between critical minimum and maximum values, where these bounds vary depending on which fiscal policy instrument is used. In other words, determinacy requires restrictions on the magnitude of γ_l^q, where $q \equiv (g, c, k, n)$.[22] By contrast, the values of γ_y^q, measuring the reaction of fiscal policy instruments to the output gap, are not found to be critical to determinacy.

Nevertheless, although not important for determinacy, we set bounds on the feedback reaction of the capital tax rate to the output gap, γ_y^k, so that the implied capital tax rate is within $0.11 < \tau_t^k < 0.51$ in the transition. In other words, since the average value of τ_t^k in the data is 0.31, we limit attention to changes in τ_t^k that are less than minus/plus 20 percentage points than in the data. We set these bounds because if the capital tax rate were left unrestricted, the policymaker would find it optimal to set it at a very low value, as one would expect given the Chamley-Judd logic. Our bounds exclude this possibility. Note that such practice is usual in the related literature; for instance, when they compute optimized rules, Schmitt-Grohé and Uribe (2007) and Kliem and Kriwoluzky (2014) impose similar intervals for monetary and fiscal policy, respectively.

5.3 Optimal Fiscal Policy Mix

We can now study optimal policy mixes when we depart from the status quo steady state and travel towards the reformed steady state

[22]For example, when we use one fiscal instrument at a time, the ranges of fiscal reaction to public liabilities are $0.027 < \gamma_l^g < 2.72$, $0.048 < \gamma_l^c < 4.34$, $0.064 < \gamma_l^k < 2.96$, and $0.063 < \gamma_l^n < 1.56$ for s_t^g, τ_t^c, τ_t^k, and τ_t^n, respectively. When we switch on all fiscal instruments, these ranges become narrower.

Table 4. Optimal Fiscal Reaction to Debt and Output with Debt Consolidation

Fiscal Instruments	Optimal Reaction to Debt	Optimal Reaction to Output
s_t^g	$\gamma_l^g = 0.5009$	$\gamma_y^g = 0$
τ_t^c	$\gamma_l^c = 0$	$\gamma_y^c = 0$
τ_t^k	$\gamma_l^k = 0.0031$	$\gamma_y^k = 2.2569$
τ_t^n	$\gamma_l^n = 0.0753$	$\gamma_y^n = 2.1360$

Notes: (i) τ_t^k is the residual instrument at the reformed steady state. (ii) At all t, $R_t \geq 1$, $0 < s_t^g, \tau_t^c, \tau_t^k, \tau_t^n < 1$. (iii) Restriction on γ_y^k so as $0.11 < \tau_t^k < 0.51$. (iv) Lifetime utility $V_0 = 79.9864$.

as defined in subsection 4.1 above. Policymakers are allowed to use different instruments in the transition and in the steady state.[23] In particular, in the reformed steady state, given the evidence in table 3, we assume that it is the capital tax rate that takes advantage of the fiscal space created once the debt burden has been reduced and sovereign premia have been eliminated; as said, a cut in the capital tax rate is the most efficient way of using this fiscal space. On the other hand, in the transition to this reformed steady state, all available fiscal instruments, and at the same time, are allowed to be used, as in the policy rules (16)–(19). Since feedback policy coefficients are chosen optimally, this will also tell us how to assign different policy instruments to different macroeconomic indicators. Results for the optimal mix during the transition to the reformed steady state are reported in table 4.

The values of the optimized feedback policy coefficients in table 4 imply a clear-cut assignment of instruments to targets. Government spending should be used to address the public debt problem, while income (capital and labor) taxes should be used to address the output problem. On the other hand, it is better to avoid changes in consumption tax rates (the optimized feedback coefficients in the

[23] Results in the special case in which the fiscal authority is restricted to use one fiscal instrument at a time are in appendix 5. These results can help to undestand the working of the model. Here, we only present the optimal mix, where instruments can be used simultaneously, to save on space.

rule for the consumption tax rate are practically zero in this case). These signs and magnitudes of the feedback policy coefficients mean that government spending should be reduced in order to bring public debt down, while, at the same time, the capital and labor tax rates should also be cut so as to stimulate the real economy in an attempt to increase the denominator in the debt-to-output ratio (impulse response functions are shown below).

Table 5 also reports some associated statistics (like elasticities and min/max values).[24] As can be seen, in the short term, public spending should fall by a lot vis-à-vis its data value so as to bring public debt down, and, at the same time, capital and labor tax rates should also fall by a lot vis-à-vis the data to stimulate the real economy. Then, over time, they return to their data average values (this is also shown by impulse response functions below). In subsection 5.6 below, we also report results when the use of policy instruments is restricted.

5.4 Welfare over Time with, and without, Debt Consolidation

Setting the feedback policy coefficients as in table 4, the associated expected discounted utility over various time horizons is reported in the first row of table 6. Studying what happens to welfare over various time horizons can be useful because, for several (e.g., political-economy) reasons, economic agents can be short-sighted. It can also help us to understand the possible conflicts between short-, medium-, and long-term effects from debt consolidation. The second row in table 6 reports results without debt consolidation, other things being equal. Thus, we again compute the best policy mix, meaning that all fiscal policy instruments at the same time are allowed to react to debt and output gaps but now the debt and output targets in the policy rules remain as in the status quo solution (see subsection 4.3 above). Finally, the last row in table 6 gives the welfare gain, or loss, of debt consolidation expressed in permanent consumption equivalent units. A positive number means that welfare would increase with debt consolidation, and vice versa: a negative number means that welfare would decrease with debt consolidation.

---

[24]The elasticities report the percentage change in the fiscal instrument with respect to a 1 percent change in the macroeconomic indicator, other things equal.

Table 5. Statistics Implied by Table 4

	Elasticity to Liabilities	Elasticity to Output	Min/Max	Five Periods Average	Ten Periods Average	Twenty Periods Average	Data Average
s_t^g	−2.52%	0%	0.1073/0.2352	0.1204	0.1383	0.1694	0.2222
τ_t^c	0%	0%	0.1756/0.1756	0.1756	0.1756	0.1756	0.1756
τ_t^k	0.01%	5.13%	0.1443/0.3118	0.1901	0.2165	0.2470	0.3118
τ_t^n	0.2%	3.61%	0.2956/0.4379	0.3383	0.3607	0.3851	0.421

Notes: See notes in table 4.

Table 6. Welfare over Different Time Horizons with, and without, Debt Consolidation

	Two Periods	Four Periods	Ten Periods	Twenty Periods	Thirty Periods
With Consolidation	1.3576	2.8154	7.4589	15.1321	22.1331
Without Consolidation[a]	(1.6251)	(3.1791)	(7.4481)	(13.4127)	(18.1904)
Welfare Gain/Loss	−0.0862	−0.0717	0.001	0.0959	0.1310

[a]The values of the optimized feedback policy coefficients without debt consolidation are $\gamma_l^g = 0.2622$, $\gamma_l^c = 0.3303$, $\gamma_l^k = 0.5819$, $\gamma_l^n = 0.2510$, $\gamma_y^g = 0.3157$, $\gamma_y^c = 0.5772$, $\gamma_y^k = 0$, $\gamma_y^n = 0.0162$.
Notes: See notes in table 4.

The results in table 6 reveal that, other things being equal, debt consolidation improves welfare only if we are relatively far-sighted. In particular, expected discounted utility is higher with debt consolidation only when we care beyond the first ten years. Reversing the argument, debt consolidation comes at a short-term cost.[25] Once the short-term pain is over, the welfare gain in consumption equivalents is substantial.[26]

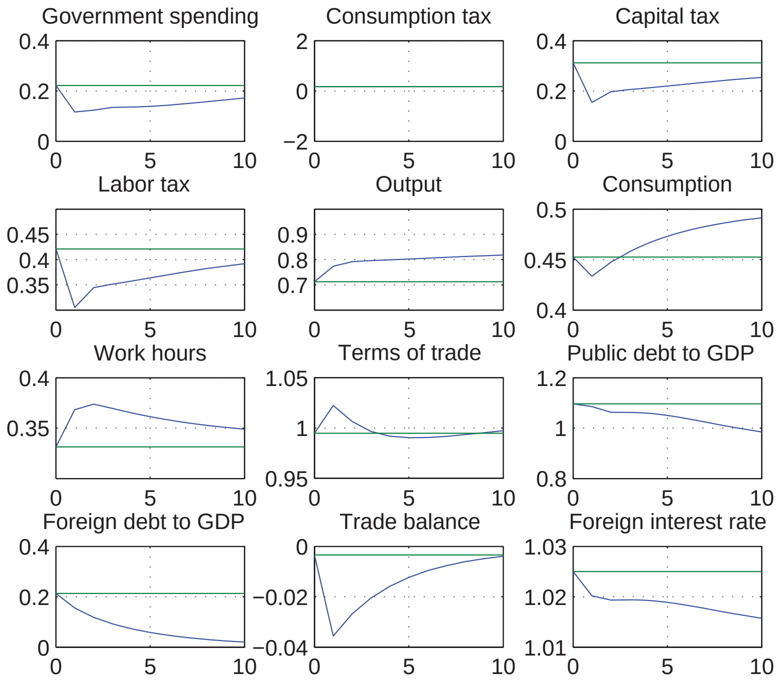
5.5 Impulse Response Functions with Debt Consolidation

To get a clearer picture of the above results, we also present the implied impulse response functions illustrating the time paths of the fiscal policy instruments used for consolidation, as well as some key macroeconomic variables. They are shown in figure 1. As can be seen, public spending should fall, while, at the same time, capital

[25]It should be pointed out that the rise in welfare is partly driven by the fact that debt consolidation and elimination of sovereign premia in the reformed long-run equilibrium allow a higher value of the time preference rate than in the pre-reformed long-run solution in section 3 (in particular, the calibrated value of β was 0.978 in the status quo steady state in section 3, while it is 0.9886 without premia). We report that the main results do not change when we allow for persistence in the change of the time preference rate as it rises from 0.978 to 0.9886. Actually, when we allow the related autoregressive parameter to be optimally chosen, along the feedback policy coefficients, its optimal value is close to zero, meaning that it is better to adopt as soon as possible the higher value of the time preference rate. Results are available upon request.

[26]Prescott (2002) finds welfare gains of similar magnitude when Japan or France adopt the tax policy or the production efficiency of the United States.

Figure 1. IRFs under Debt Consolidation with the Optimal Fiscal Mix



Note: IRFs are in levels and converge to the reformed steady state, while the solid horizontal line indicates the point of departure (status quo value).

and labor tax rates should be cut, for the reasons discussed above. This optimal mix allows a gradual reduction in the ratio of public debt to GDP, in the ratio of foreign debt to GDP, and in the interest rate premium. Private consumption falls in the short term, as a consequence of debt consolidation, but recovers soon. Hours of work need to rise (or leisure to fall) for some time. All variables converge to their new, reformed values over time.

5.6 Further Restrictions on the Use of Fiscal Policy Instruments

One could argue that the values of tax–spending policy instruments cannot differ substantially from those in the historical data (for various political-economy reasons). Therefore, we now redo the main computations, restricting the magnitude of feedback coefficients in the policy rules so that all tax–spending policy instruments cannot change by more than, say, 10 percentage points from their averages in the data. The new results are reported in appendix 5. Although, obviously, feedback policy coefficients are now smaller, the best fiscal policy mix again implies that we should earmark public spending for the reduction of public debt and, at the same time, cut taxes to mitigate the recessionary effects of debt consolidation. The only difference is that now, since cuts in income (capital and labor) taxes are restricted, we should also cut consumption taxes.

In the same appendix (appendix 5), we compare results when the fiscal authorities use the optimal fiscal mix with results when they are restricted to using one fiscal instrument at a time only. The message from the new impulse response functions is that the reduction in public debt is more gradual when we use all fiscal instruments, and this allows a smaller fall in private consumption, than when the fiscal authorities are restricted to using one instrument at a time only. This is intuitive: a policy mix gives more choices.

6. Sensitivity Analysis

This section checks the sensitivity of the above results. We start with changes in parameter values and then study robustness to more substantial modeling changes. To save on space, we will selectively provide some results only (a full set of results is available upon request from the authors).

6.1 Changes in Parameter Values

We start with the value of the public debt threshold parameter, $\overline{d}$, in the interest rate premium equation (15). Recall that so far we have set $\overline{d} = 0.9$. Our qualitative results do not depend on this value. For instance, in appendix 6, we present the main results with $\overline{d} = 0.8$ and $\overline{d} = 1$. In this case, as said above, we need to recalibrate the value of ψ so as to hit the data again; the new values are, respectively, $\psi = 0.0319$ and $\psi = 0.108$.

Our results are also robust to changes in the assumed value of net foreign debt in the steady state of the reformed economy. Recall that so far we have solved for the reformed economy assuming a zero net foreign debt position at steady state, $\widetilde{f} = 0$. Our main results remain unchanged when we instead set $\widetilde{f} = 0.1$ and $\widetilde{f} = 0.2109$ in the reformed steady state (where 0.2109 is the average value of the country's foreign debt in the data). Results for these two cases are presented in appendix 7.

Our qualitative results are also robust to changes in other model parameters. For instance, we have experimented with changes in the values of the Calvo parameter in the firm's problem, θ, and the adjustment cost parameters on foreign private assets/debt, ϕ^h; foreign public debt, ϕ^g; and physical capital, ξ. We report that our main results do not change when we set θ at, say, 0, 0.1, or 0.9 (we also report that as price stickiness, θ, rises, the optimal fiscal reaction to public debt becomes milder), when $0.2 \leq \phi^h \leq 0.5$, and when $0 \leq \xi \leq 2$, while the value of ϕ^g has not been found to be important. Similarly, our results remain unchanged for $0.8 \leq \gamma \leq 0.99$, which measures the sensitivity of exports to changes in the terms of trade in equation (20). Also, our main results do not depend on the value of χ_g, namely, how much agents value public consumption spending in the utility function. The values of the labor supply parameters, χ_n and η, are not crucial either; nevertheless, we report that when $\eta = 0.5$, the reaction of the labor tax rate to the debt target is zero, i.e., $\gamma_l^n = 0$, while when $\eta = 2$, we get $\gamma_l^n = 0.3761$. That is, as η (namely, the inverse of the Frisch elasticity) rises, the reaction of the labor tax rate to the debt target becomes stronger, which, in turn, means that the net cut in the labor tax rate should be smaller. As said, the above results are available upon request.

6.2 Changes in Policy Variables

Our results are also robust to the specific way we model the fiscal policy instruments. For instance, the main results remain unaffected when we allow for persistence in the feedback policy rules, (16)–(19), in the sense that—for example, in the case of the consumption tax rate—we use

$$\tau_t^c = (1 - \rho^{\tau^c})\tau^c + \rho^{\tau^c}\tau_{t-1}^c + \gamma_l^c \left(l_{t-1} - l\right) + \gamma_y^c \left(y_t^H - y^H\right), \quad (22)$$

where $0 \leq \rho^{\tau^c} \leq 1$ is an autoregressive policy parameter and the initial value of the policy instrument is its data average value as reported in tables 1 and 2. Actually, we have also allowed ρ^{τ^c} to be optimally chosen, jointly with the other feedback policy coefficients, and the main results again do not change. Interestingly, the optimized values of these autoregressive policy parameters are found to be relatively small, meaning that it is better to adjust the policy instrument(s) relatively soon.

Also, following several related papers (see, e.g., Coenen, Mohr, and Straub 2008, Forni, Gerali, and Pisani 2010b, and Erceg and Lindé 2013), we have experimented with time-varying debt policy targets. Thus, instead of using a constant-over-time debt policy target, l, as in equations (16)–(19) above, we assume that the target, defined now as l_t^*, follows an $AR(1)$ process of the form

$$l_t^* = \left(1 - \rho^l\right) l + \rho^l l_{t-1}^*, \tag{23}$$

where $0 \leq \rho^l \leq 1$ is an autogressive policy parameter and the initial value of the target is given by its data average value in tables 1 and 2. We report that our main results remain the same under this new specification.

Our results are also robust to adding more macroeconomic indicators in the feedback policy rules (like inflation or terms of trade). We have also experimented with changes in some exogenous policy instruments, which have been kept constant so far, like the fraction of public debt held by domestic agents relative to foreign investors, λ. The latter has so far been kept constant and equal to its average value on the data, 0.64. When we experiment with $\lambda = 0.54$ and $\lambda = 0.74$, the results do not change.

6.3 Allowing for New Shocks

Our results are also robust to allowing for a more volatile economy. This can be captured by increasing the standard deviation of the existing TFP shock and/or by adding new shocks. Specifically, regarding new shocks, we have experimented with adding shocks to the fiscal policy rules in subsection 2.7, to the time-varying debt policy target presented in subsection 6.2 above, or to the world interest rate in equation (15). The main results again do not change. In particular, regarding shocks to the world interest rate, and following the specification of García-Cicco, Pancrazi, and Uribe (2010), we augment equation (15) by

$$Q_t = Q_t^* + \psi \left(e^{\left(\frac{D_t}{P_t^H Y_t^H} - \overline{d} \right)} - 1 \right) + \left(e^{\varepsilon_t^{\mu_t^q - 1}} - 1 \right), \tag{24}$$

where

$$\log \left(\mu_t^q \right) = \rho^q \log \left(\mu_{t-1}^q \right) + \varepsilon_t^q, \tag{25}$$

where $0 \leq \rho^q \leq 1$ is a parameter and ε_t^q is an iid shock. In our experiments, we set 0.9845 for ρ^q and 0.0487 for the standard deviation of ε_t^q.[27] The new results are reported in appendix 8. As can be seen, the main messages remain the same. This is not suprising: the key driver of transition dynamics is policy reforms, such as debt consolidation, rather than cyclical fluctuations generated by exogenous shocks.

6.4 Transition to the Ramsey Steady State

So far, we have studied the optimal transition of the economy from the status quo to an arbitrarily reformed steady state. A potential criticism (see, e.g., Schmitt-Grohé and Uribe 2009, chapter 3) might be that the accuracy of approximation (in our paper, first and second order) is poor because when the policymaker chooses the feedback policy coefficients in such cases, he/she may want to use his/her choices in order to influence the mean of variables, rather than to stabilize the economy around the assumed steady state. A way of checking whether this affects the main results is to approximate the economy around the Ramsey steady state, meaning that the steady-state policy targets in the feedback rules (16)–(19) are the Ramsey steady-state values of the corresponding variables.

Therefore, this subsection examines the optimal transition of the economy from the status quo to the Ramsey steady state and compares this with the case studied so far, in which we computed the optimal transition to an arbitrarily reformed steady state. By optimal transition, we again mean that we select the feedback policy coefficients so as to maximize the household's welfare criterion as

[27] These estimates follow from OLS regressions of (25) where μ_t^q denotes the deviation of the ten-year German bond rate from its average value on the data. The data for the ten-year German bond are from Eurostat, 2001–13.

above. The difference is that now the steady state, around which we approximate, is the steady state of the Ramsey second-best policy problem.[28]

The first step is to solve the Ramsey second-best policy problem. In this problem, when we use the so-called dual approach, the government chooses the time paths of all endogenous variables of the DE system plus the fiscal policy instruments taken as given at the DE level (see, e.g., Schmitt-Grohé and Uribe 2006 for a related problem). For simplicity, and following the same authors, we solve for the so-called timeless Ramsey equilibrium, which means that the optimality conditions in the initial periods do not differ from those in later periods. Besides, to make the comparison with the previous cases meaningful, we again assume, as we have done so far in the steady state of the arbitrarily reformed economy, that, in the steady state of the Ramsey problem, the ratio of public debt to GDP is set at 0.9 and that the country's net foreign debt position is zero.[29] Details of the Ramsey policy problem are in appendix 9.

When we solve this Ramsey problem numerically using Dynare, the resulting steady-state solution (reported in appendix 9) gives us an unrealistically high consumption tax rate and an equally large labor subsidy, i.e., $\tau^n < 0$. This is a well-known property of the Ramsey problem, when the policymaker also has access to consumption taxes (see the early papers by Lansing 1999, Coleman 2000, and many others since then).[30] In addition, our DE economy does not seem to converge to this unrestricted Ramsey steady state. To overcome these problems, and given the difficulty of setting negative labor tax rates in reality, we solve the Ramsey problem by not allowing the government to choose the labor tax rate, which is now constrained to remain fixed at, say, zero ($\tau^n = 0$). Actually, most of

---

[28]Schmitt-Grohé and Uribe (2006, 2007) and Kliem and Kriwoluzky (2014), on the other hand, select the feedback policy coefficients so as to minimize the distance from the Ramsey solution.

[29]Actually, setting some debt values at exogenous values is necessary to get a well-defined steady-state system in the Ramsey problem. See also, e.g., Schmitt-Grohé and Uribe (2004a). We provide details in appendix 9.

[30]Another related well-known property is that, at least in the baseline neoclassical growth model without market frictions, the Ramsey government can implement the first-best allocation if the policymaker has access to consumption, labor, and capital taxes.

Table 7. Ramsey Steady-State Solution with $\tau^n = 0$

Variables	Description	Steady-State Solution
u	Period Utility	1.0950
y^H	Output	1.1162
TT	Terms of Trade	1.01
τ^k	"Optimal" Capital Tax Rate	0.0500
τ^c	"Optimal" Consumption Tax Rate	0.810199
$Q - Q^*$	Interest Rate Premium	0
$TT^{1-\nu}\frac{c}{y^H}$	Consumption as Share of GDP	0.5443
$\frac{k}{y^H}$	Physical Capital as Share of GDP	5.8387
$\frac{TT^{\nu^*} f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.3240
$\frac{TT^{1-\nu} d}{y^H}$	Total Public Debt as Share of GDP	0.9
$\widetilde{f}$	Total Foreign Debt as Share of GDP	0

the above-mentioned literature on Ramsey policy with consumption taxes works similarly. The steady-state solution of this "restricted" Ramsey problem is reported in table 7.

As can be seen, the implied utility and output are higher in the Ramsey steady state than in all other cases studied so far. This is as expected. It is also worth noticing that, in this solution, the capital tax rate is very low (but not zero since there are several market imperfections in the model) and the consumption tax rate is well defined (although high, as one would expect for the reasons discussed above).

The next step is to compute the optimal transition from the status quo steady state in table 2 to the Ramsey steady state in table 7. Results for the optimal fiscal policy mix and the associated welfare over different time horizons are reported in tables 8 and 9, respectively. These two tables need to be compared with tables 4 and 6, respectively. As can be seen, the main results are not affected. Namely, it is again optimal to earmark public spending for the reduction of public debt and, at the same time, to use the tax rates for the stimulation of the real economy. The difference is that since changes in the labor tax are now restricted (the Ramsey labor tax rate has

Table 8. Optimal Fiscal Reaction to Debt and Output with Debt Consolidation and a Ramsey Steady State

Fiscal Instruments	Optimal Reaction to Debt	Optimal Reaction to Output
s_t^g	$\gamma_l^g = 0.6515$	$\gamma_y^g = 0.0102$
τ_t^c	$\gamma_l^c = 0$	$\gamma_y^c = 1.9621$
τ_t^k	$\gamma_l^k = 0$	$\gamma_y^k = 0.6183$
τ_t^n	$\gamma_l^n = 0.0086$	$\gamma_y^n = 0.0002$

Notes: See notes (i)–(iii) in table 4. (iv) $V_0 = 92.4789$. (v) Optimal degree of persistence in fiscal instruments, $\rho = 0.2533$.

Table 9. Welfare over Different Time Horizons with, and without, Debt Consolidation and a Ramsey Steady State

	Two Periods	Four Periods	Ten Periods	Twenty Periods	Thirty Periods
With Consolidation	1.7351	3.3455	8.2102	16.6584	24.6462
Without Consolidation[a]	(1.6251)	(3.1791)	(7.4481)	(13.4127)	(18.1904)
Welfare Gain/Loss	0.0378	0.0346	0.0761	0.1888	0.2789

[a]The values of the optimized feedback policy coefficients without debt consolidation are as in table 6.

Notes: See notes in table 8.

been set to zero), it is also optimal to make use of the consumption tax rate. Namely, it is optimal, in the short term, to cut not only the capital tax rate but also the consumption tax rate; in later periods, the consumption tax rate rises substantially, converging to its high Ramsey steady-state value in table 7. Observe that now, as shown in table 9, the case with debt consolidation is superior to the reference case without debt consolidation over all time horizons, even in the short term. Thus, an intuitive general message is that a proper use of fiscal policy can mitigate (in our case, avoid) the short-term pain from debt consolidation. But, of course, as is typically the case, the optimal Ramsey values for the fiscal policy instruments are far away from their values in the data.

Table 10. Optimal Fiscal Reaction to Debt and Output with Debt Consolidation and when the Premium Depends on Foreign Debt

Fiscal Instruments	Optimal Reaction to Debt	Optimal Reaction to Output
s_t^g	$\gamma_l^g = 0.5322$	$\gamma_y^g = 0.0005$
τ_t^c	$\gamma_l^c = 0$	$\gamma_y^c = 0.2912$
τ_t^k	$\gamma_l^k = 0$	$\gamma_y^k = 2.9810$
τ_t^n	$\gamma_l^n = 0.283$	$\gamma_y^n = 2.9578$

Notes: See notes (i)–(iii) as in table 4. (iv) $V_0 = 79.9909$.

6.5 The Interest Rate Premium as a Function of Net Foreign Debt

So far, we have assumed that the sovereign interest rate premium in equation (15) is a function of the ratio of public debt to GDP. We now assume that the premium is a function of the country's net foreign debt ratio, $\widetilde{f} \equiv \frac{(1-\lambda)TT^{1-\nu}d - TT^{\nu^*}f^h}{y^H}$. The latter is another obvious candidate for the emergence of country interest rate premia. In particular, equation (15) changes to (assuming a zero threshold parameter in this case)[31]

$$Q_t - Q^* = \psi\left(e^{\frac{(1-\lambda)dTT^{1-\nu} - TT^{\nu^*}f^h}{y^H}} - 1\right). \qquad (26)$$

The new results are summarized in tables 10 and 11. As can be seen by comparing tables 10 and 11 with tables 4 and 6, respectively, the main messages remain the same.

7. What Would Have Happened under Flexible Exchange Rates?

This section resolves the baseline model developed in section 2 under the fiction of flexible exchange rates, other things being equal. Again,

[31]The value of ψ is recalibrated in the same way as explained in subsections 3.1 and 6.1 above (namely, to hit the net foreign debt position in the data).

Table 11. Welfare over Different Time Horizons with, and without, Debt Consolidation and when the Premium Depends on Foreign Debt

	Two Periods	Four Periods	Ten Periods	Twenty Periods	Thirty Periods
With Consolidation	1.3667	2.8256	7.4659	15.1503	22.1528
Without Consolidation[a]	(1.6251)	(3.1793)	(7.4468)	(13.4065)	(18.1785)
Welfare Gain/Loss	−0.0834	−0.0698	0.0018	0.0974	0.1635

[a]In the reference case without debt consolidation, the optimal feedbacks are $\gamma_l^g = 0.3303$, $\gamma_l^c = 0.0033$, $\gamma_l^k = 0.7693$, $\gamma_l^n = 0.2089$, $\gamma_y^g = 0.3042$, $\gamma_y^c = 0.4519$, $\gamma_y^k = 0.0041$, $\gamma_y^n = 0.0045$.
Notes: See notes in table 10.

the initial conditions for the state variables will be those of the steady-state solution of the status quo model. In terms of modeling, the only difference from the model in section 2 is that now the exchange rate becomes an endogenous variable. Thus, R_t and S_t exchange places. The former was endogenous in section 2, while now it is the latter that becomes endogenous, with the former being free to follow a national Taylor-type rule for the nominal interest rate. In particular, we postulate

$$\log\left(\frac{R_t}{R}\right) = \phi_\pi \log\left(\frac{\Pi_t}{\Pi}\right) + \phi_y \log\left(\frac{y_t^H}{y^H}\right) + \phi_\epsilon \log\left(\frac{\epsilon_t}{\epsilon}\right), \qquad (27)$$

where ϕ_π, ϕ_y, $\phi_\epsilon \geq 0$ are feedback monetary policy coefficients on price inflation, output, and exchange rate depreciation, respectively, as deviations from their steady-state values. This is like the fiscal policy rules above. All feedback (monetary and fiscal) policy coefficients are again computed optimally, as described in section 4.

Since money is neutral in the long run (and $\Pi^{H*} \equiv 1$), a switch to flexible exchange rates does not affect the solution of real variables in the steady state. Any differences will thus arise in the transition only, during which money is not neutral because of Calvo-type nominal fixities. Results are reported in tables 12 and 13, which need to be compared with tables 4 and 6, respectively (appendix 10 also presents impulse response functions). The computed values of ϕ_π and ϕ_y are as typically found in the related literature (see, e.g.,

Table 12. Optimal Reaction to Inflation, Depreciation, Debt, and Output with Debt Consolidation (Optimal Monetary and Fiscal Policy Mix) under Flexible Exchange Rates

Monetary Policy		Fiscal Policy		
Monetary Instrument	**Optimal Reaction to Inflation, Output, and Depreciation**	**Fiscal Instruments**	**Optimal Reaction to Debt**	**Optimal Reaction to Output**
R_t	$\phi_\pi = 1.36$ $\phi_y = 0$ $\phi_\epsilon = 100$	s_t^g τ_t^c τ_t^k τ_t^n	$\gamma_l^g = 0.5268$ $\gamma_l^c = 0$ $\gamma_l^k = 0.0012$ $\gamma_l^n = 0.0744$	$\gamma_y^g = 0.0001$ $\gamma_y^c = 0$ $\gamma_y^k = 2.782$ $\gamma_y^n = 2.1876$

Notes: See notes (i)–(iii) in table 4. (iv) $V_0 = 79.9954$.

Table 13. Welfare over Different Time Horizons with, and without, Debt Consolidation under Flexible Exchange Rates

	Two Periods	Four Periods	Ten Periods	Twenty Periods	Thirty Periods
With Consolidation	1.3482	2.7976	7.4448	15.1411	22.1486
Without Consolidation[a]	(1.6247)	(3.1786)	(7.4451)	(13.4043)	(18.1768)
Welfare Gain/Loss	−0.0890	−0.0750	−0.00005	0.0969	0.1634

[a]In the case without debt consolidation, the optimal feedbacks are $\phi_\pi = 1.3972$, $\phi_y = 0$, $\phi_\epsilon = 0.0002$, $\gamma_l^g = 0.2515$, $\gamma_l^c = 0.0011$, $\gamma_l^k = 0.6417$, $\gamma_l^n = 0.3298$, $\gamma_y^g = 0.3572$, $\gamma_y^c = 0.3132$, $\gamma_y^k = 0.0169$, $\gamma_y^n = 0.0001$.
Notes: See notes in table 12.

Schmitt-Grohé and Uribe 2007), while the high value of ϕ_ϵ implies that exchange rate stabilization is desirable. But, in the context of our paper, the key result is that the optimal fiscal policy mix remains as above. Notice also that the associated welfare is only slightly higher than that without monetary policy independence in table 4. In other words, to the extent that feedback policy coefficients are selected optimally, the loss of monetary policy independence is not a big loss, at least in this class of New Keynesian models with Calvo- or Rotemberg-type nominal fixities. This is in line with the related literature (see Schmitt-Grohé and Uribe 2016).[32]

[32]Schmitt-Grohé and Uribe (2016) give the following explanation for the small differences in macro performance under fixed and flexible exchange rates in this class of models: increases in unemployment during recessions are roughly offset by rises in work hours during expansions, so that the average level of employment, and hence welfare, is affected relatively little by the exchange rate regime (we report that, in our model, impulse response functions for hours of work and real wages are very similar under flexible and fixed exchange rates). It therefore seems that one has to add extra forms of nominal fixities to make flexible exchange rates, and hence the use of independent monetary policy, more desirable. For instance, Schmitt-Grohé and Uribe (2016) add downward nominal wage rigidity in a small open-economy model and also assume a labor contract according to which employment is demand determined during recessions but demand-supply determined during booms. This implies that aggregate fluctuations cause higher unemployment on average, so that having an extra instrument for stabilization, such as independent monetary policy, becomes useful.

8. Concluding Remarks and Possible Extensions

This paper has studied fiscal policy action in a New Keynesian model of a small open economy facing debt-elastic interest rate premia and not being able to use monetary policy. Our analysis was based on optimized, simple, and implementable feedback policy rules for various categories of tax rates and public spending.

Since the main results have been listed in the introduction already, we close with some possible extensions. Here, we have focused on the macroeconomic, or aggregate, implications of alternative debt consolidation policies, leaving out the issue of distributional implications. It would be interesting to add heterogeneity both in terms of economic agents within the country and in terms of countries. For instance, within each country, we could distinguish between those who have access to financial markets and those who just work and consume (the so-called rule-of-thumb consumers) or between those working in the private sector and those working in the public sector (public employees). It would be also interesting to use a two-country model, where countries can differ in, say, fiscal imbalances and/or time preferences, and so study the asymmetric cross-border effects of national stabilization and debt consolidation policies. We leave these extensions for future work.

References

Alesina, A., C. Favero, and F. Giavazzi. 2012. "The Output Effect of Fiscal Consolidations." NBER Working Paper No. 18336.

Almeida, V., G. Castro, R. M. Félix, and J. R. Maria. 2013. "Fiscal Consolidation in a Small Euro-Area Economy." *International Journal of Central Banking* 9 (4): 1–38.

Andrès, J., and R. Doménech. 2006. "Automatic Stabilizers, Fiscal Rules and Macroeconomic Stability." *European Economic Review* 50 (6): 1487–1506.

Ardagna, S., F. Caselli, and T. Lane. 2008. "Fiscal Discipline and the Cost of Public Debt Service: Some Estimates for OECD Countries." *B.E. Journal of Macroeconomics* 7 (1): 1–35.

Arellano, C. 2008. "Default Risk and Income Fluctuations in Emerging Economies." *American Economic Review* 98 (3): 690–712.

Batini, N., G. Gallegari, and M. Giovanni. 2012. "Successful Austerity in the United States, Europe and Japan." IMF Working Paper No. 12/190.

Batini, N., P. Levine, and J. Pearlman. 2008. "Monetary and Fiscal Rules in an Emerging Small Open Economy." Working Paper No. 08/10, Centre for Dynamic Macroeconomic Analysis, University of Surrey.

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): 926–48.

Benigno, P., and M. Woodford. 2012. "Linear-Quadratic Approximation of Optimal Policy Problems." *Journal of Economic Theory* 147 (1): 1–42.

Bi, H. 2010. "Optimal Debt Targeting Rules for Small Open Economies." Mimeo.

———. 2012. "Sovereign Default Risk Premia, Fiscal Limits, and Fiscal Policy." *European Economic Review* 56 (3): 389–410.

Bi, H., and M. Kumhof. 2011. "Jointly Optimal Monetary and Fiscal Policy Rules under Liquidity Constraints." *Journal of Macroeconomics* 33 (3): 373–89.

Bi, H., E. Leeper, and C. Leith. 2013. "Uncertain Fiscal Consolidations." *Economic Journal* 123 (566): F31–F63.

Blanchard, O., and F. Giavazzi. 2003. "Macroeconomic Effects of Regulation and Deregulation in Goods and Labor Markets." *Quarterly Journal of Economics* 118 (3): 879–907.

Cantore, C., P. Levine, G. Melina, and J. Pearlman. 2012. "Optimal Fiscal and Monetary Rules in Normal and Abnormal Times." Mimeo, University of Surrey.

Checherita-Westphal, C., and P. Rother. 2012. "The Impact of High Government Debt on Economic Growth and Its Channels: An Empirical Investigation for the Euro Area." *European Economic Review* 56 (7): 1392–1405.

Christiano, L., M. Trabandt, and K. Walentin. 2011. "Introducing Financial Frictions and Unemployment into a Small Open Economy Model." *Journal of Economics Dynamics and Control* 35 (12): 1999–2041.

Coenen, G., M. Mohr, and R. Straub. 2008. "Fiscal Consolidation in the Euro Area: Long-Run Benefits and Short-Run Costs." *Economic Modelling* 25 (5): 912–32.

Cogan, J., J. Taylor, V. Wieland, and M. Wolters. 2013. "Fiscal Consolidation Strategy." *Journal of Economic Dynamics and Control* 37 (2): 404–21.

Coleman, W. J. 2000. "Welfare and Optimum Dynamic Taxation of Consumption and Income." *Journal of Public Economics* 76 (1): 1–39.

Corsetti, G., K. Kuester, A. Meier, and G. Muller. 2013. "Sovereign Risk, Fiscal Policy, and Macroeconomic Stability." *Economic Journal* 123 (566): F99–F132.

Davig, T., and E. Leeper. 2011. "Monetary-Fiscal Policy Interactions and Fiscal Stimulus." *European Economic Review* 55 (2): 211–27.

Eaton, J., and M. Gersovitz. 1981. "Debt with Potential Repudiation: Theoretical and Empirical Analysis." *Review of Economic Studies* 48 (2): 289–309.

Erceg, C., and J. Lindé. 2012. "Fiscal Consolidation in an Open Economy." *American Economic Review* 102 (3): 186–91.

———. 2013. "Fiscal Consolidation in a Currency Union: Spending Cuts vs. Tax Hikes." *Journal of Economic Dynamics and Control* 37 (2): 422–45.

European Commission. 2011. "Public Finances in EMU—2011." Report No. 3/2011, European Economy Series.

———. 2012. "Public Finances in EMU—2012." Report No. 4/2012, European Economy Series.

European Economic Advisory Group. 2014. "The EEAG Report on the European Economy." CESifo.

Forni, L., A. Gerali, and M. Pisani. 2010a. "Macroeconomic Effects of Fiscal Consolidation in a Monetary Union: The Case of Italy." Discussion Paper No. 747, Banca d'Italia.

———. 2010b. "The Macroeconomics of Fiscal Consolidation in Euro Area Countries." *Journal of Economic Dynamics and Control* 34 (9): 1791–1812.

Galí, J. 2008. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework.* Princeton University Press.

Galí, J., and T. Monacelli. 2005. "Monetary Policy and Exchange Rate Volatility in a Small Open Economy." *Review of Economic Studies* 72 (3): 707–34.

———. 2008. "Optimal Monetary and Exchange Rate Policy in a Currency Union." *Journal of International Economics* 76 (1): 116–32.

Garcia-Cicco, J., R. Pancrazi, and M. Uribe. 2010. "Real Business Cycles in Emerging Economies?" *American Economic Review* 100 (5): 2510–31.

Herz, B., and S. Hohberger. 2013. "Fiscal Policy, Monetary Regimes and Current Account Dynamics." *Review of International Economics* 21 (1): 118–36.

Kirsanova, T., C. Leith, and S. Wren-Lewis. 2009. "Monetary and Fiscal Policy Interaction: The Current Consensus Assignment in the Light of Recent Developments." *Economic Journal* 119 (541): F482–F495.

Kirsanova, T., M. Satchi, D. Vines, and S. Wren-Lewis. 2007. "Optimal Fiscal Policy Rules in a Monetary Union." *Journal of Money, Credit and Banking* 39 (7): 1759–84.

Kirsanova, T., and S. Wren-Lewis. 2012. "Optimal Fiscal Feedback on Debt in an Economy with Nominal Rigidities." *Economic Journal* 122 (559): 238–64.

Kliem, M., and A. Kriwoluzky. 2014. "Toward a Taylor Rule for Fiscal Policy." *Review of Economic Dynamics* 17 (2): 294–302.

Lansing, K. 1999. "Optimal Redistributive Capital Taxation in a Neoclassical Growth Model." *Journal of Public Economics* 73 (3): 423–53.

Leeper, E., M. Plante, and N. Traum. 2009. "Dynamics of Fiscal Financing in the United States." *Journal of Econometrics* 156 (2): 304–21.

Leith, C., and S. Wren-Lewis. 2008. "Interactions between Monetary and Fiscal Policy under Flexible Exchange Rates." *Journal of Economic Dynamics and Control* 32 (9): 2854–82.

Mendoza, E., and L. Tesar. 2005. "Why Hasn't Tax Competition Triggered a Race to the Bottom? Some Quantitative Lessons from the EU." *Journal of Monetary Economics* 52 (1): 163–204.

Pappa, E., R. Sajedi, and E. Vella. 2015. "Fiscal Consolidation with Tax Evasion and Corruption." *Journal of International Economics* 96 (S1): S56–S75.

Pappa, E., and V. Vassilatos. 2007. “The Unbearable Tightness of Being in a Monetary Union: Fiscal Restrictions and Regional Stability.” *European Economic Review* 51 (6): 1492–1513.

Perotti, R. 1996. “Fiscal Consolidation in Europe: Composition Matters.” *American Economic Review* 86 (2): 105–10.

Philippopoulos, A., P. Varthalitis, and V. Vassilatos. 2015. “Optimal Fiscal and Monetary Policy Action in a Closed Economy.” *Economic Modelling* 48 (August): 175–88.

Prescott, E. 2002. “Prosperity and Depression.” *American Economic Review* 92 (2): 1–15.

Reinhart, C., and K. Rogoff. 2010. “Growth in a Time of Debt.” *American Economic Review* 100 (2): 573–78.

Roeger, W., and J. in ’t Veld. 2013. “Expected Sovereign Defaults and Fiscal Consolidations.” Economic Paper No. 479, European Commission.

Schmitt-Grohé, S., and M. Uribe. 2003. “Closing Small Open Economies.” *Journal of International Economics* 61 (1): 163–85.

———. 2004a. “Optimal Fiscal and Monetary Policy under Sticky Prices.” *Journal of Economic Theory* 114 (2): 198–230.

———. 2004b. “Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function.” *Journal of Economic Dynamics and Control* 28 (4): 755–75.

———. 2006. “Optimal Fiscal and Monetary Policy in a Medium-Scale Macroeconomic Model.” In *NBER Macroeconomics Annual*, ed. M. Gertler and K. Rogoff, 385–425. Cambridge, MA: MIT Press.

———. 2007. “Optimal Simple and Implementable Monetary and Fiscal Rules.” *Journal of Monetary Economics* 54 (6): 1702–25.

———. 2009. “Perturbation Methods for the Numerical Analysis of DSGE Models: Lecture Notes.” Mimeo, Columbia University.

———. 2016. “Downward Nominal Wage Rigidity, Currency Pegs, and Involuntary Unemployment.” *Journal of Political Economy* 124 (5): 1466–1514.

Wren-Lewis, S. 2010. “Macroeconomic Policy in Light of the Credit Crunch: The Return of Counter-Cyclical Fiscal Policy?” *Oxford Review of Economic Policy* 26 (1): 71–86.

Online Appendix to Fiscal Consolidation in an Open Economy with Sovereign Premia and without Monetary Policy Independence*

Apostolis Philippopoulos,[a,b] Petros Varthalitis,[c] and Vanghelis Vassilatos[a]
[a]Athens University of Economics and Business
[b]CESifo
[c]Economic and Social Research Institute, Trinity College Dublin

Appendix 1. Households

This appendix presents and solves the problem of the household in some detail. There are $i = 1, 2, \ldots, N$ identical domestic households that act competitively.

Household's Problem

Each household i maximizes expected lifetime utility given by

$$E_0 \sum_{t=0}^{\infty} \beta^t U\left(c_{i,t}, n_{i,t}, g_t\right), \tag{1}$$

where $c_{i,t}$ is i's consumption bundle, $n_{i,t}$ is i's hours of work, g_t is per capita public spending, $0 < \beta < 1$ is the time preference rate, and E_0 is the rational expectations operator conditional on the information set.

The consumption bundle, $c_{i,t}$, is defined as

$$c_{i,t} = \frac{\left(c_{i,t}^{H}\right)^{\nu}\left(c_{i,t}^{F}\right)^{1-\nu}}{\nu^{\nu}(1-\nu)^{1-\nu}}, \tag{2}$$

*Corresponding author: Apostolis Philippopoulos, Department of Economics, Athens University of Economics and Business, Athens 10434, Greece. Tel: +30-210-8203357. E-mail: aphil@aueb.gr.

where $c_{i,t}^{H}$ and $c_{i,t}^{F}$ denote the composite domestic good and the composite foreign good, respectively, and ν is the degree of preference for the domestic good.

We define the two composites, $c_{i,t}^{H}$ and $c_{i,t}^{F}$, as

$$c_{i,t}^{H} = \left[\sum_{h=1}^{N}[c_{i,t}^{H}(h)]^{\frac{\phi-1}{\phi}}\right]^{\frac{\phi}{\phi-1}} \tag{3}$$

$$c_{i,t}^{F} = \left[\sum_{f=1}^{N}[c_{i,t}^{F}(f)]^{\frac{\phi-1}{\phi}}\right]^{\frac{\phi}{\phi-1}}, \tag{4}$$

where $\phi > 0$ is the elasticity of substitution across goods produced in the domestic country.

The period budget constraint of each i expressed in real terms is (see, e.g., Benigno and Thoenissen 2008, for similar modeling)

$$\begin{aligned}
&(1+\tau_t^c)\left[\frac{P_t^H}{P_t}c_{i,t}^H + \frac{P_t^F}{P_t}c_{i,t}^F\right] + \frac{P_t^H}{P_t}x_{i,t} + b_{i,t} + \frac{S_t P_t^*}{P_t}f_{i,t}^h \\
&\quad + \frac{\phi^h}{2}\left(\frac{S_t P_t^*}{P_t}f_{i,t}^h - \frac{SP^*}{P}f^h\right)^2 \\
&\quad = \left(1-\tau_t^k\right)\left[r_t^k\frac{P_t^H}{P_t}k_{i,t-1} + \widetilde{\omega}_{i,t}\right] + (1-\tau_t^n)\,w_t n_{i,t} \\
&\qquad + R_{t-1}\frac{P_{t-1}}{P_t}b_{i,t-1} + Q_{t-1}\frac{S_t P_t^*}{P_t}\frac{P_{t-1}^*}{P_t^*}f_{i,t-1}^h - \tau_{i,t}^l,
\end{aligned} \tag{5}$$

where P_t is the consumer price index (CPI), P_t^H is the price index of home tradables, P_t^F is the price index of foreign tradables (expressed in domestic currency), $x_{i,t}$ is i's domestic investment, $b_{i,t}$ is i's end-of-period real domestic government bonds, $f_{i,t}^h$ is i's end-of-period real internationally traded assets denominated in foreign currency, r_t^k is the real return to inherited domestic capital, $k_{i,t-1}$, $\widetilde{\omega}_{i,t}$ is i's real dividends received by domestic firms, w_t is the real wage rate, $R_{t-1} \geq 1$ is the gross nominal return to domestic government bonds between $t-1$ and t, $Q_{t-1} \geq 1$ is the gross nominal return to international assets between $t-1$ and t, $\tau_{i,t}^l$ are real lump-sum taxes if positive (or transfers if negative) to each household, and

τ_t^c, τ_t^k, τ_t^n are tax rates on consumption, capital income, and labor income, respectively. The parameter $\phi^h \geq 0$ captures adjustment costs related to private foreign assets, where variables without time subscripts denote steady-state values.

Each household i also faces the budget constraints

$$P_t c_{i,t} = P_t^H c_{i,t}^H + P_t^F c_{i,t}^F \tag{6}$$

$$P_t^H c_{i,t}^H = \sum_{h=1}^{N} P_t^H(h) c_{i,t}^H(h) \tag{7}$$

$$P_t^F c_{i,t}^F = \sum_{f=1}^{N} P_t^F(f) c_{i,t}^F(f), \tag{8}$$

where $P_t^H(h)$ is the price of each variety h produced at home and $P_t^F(f)$ is the price of each variety f produced abroad (expressed in domestic currency).

Finally, the law of motion of physical capital for household i is

$$k_{i,t} = (1-\delta)k_{i,t-1} + x_{i,t} - \frac{\xi}{2}\left(\frac{k_{i,t}}{k_{i,t-1}} - 1\right)^2 k_{i,t-1}, \tag{9}$$

where $0 < \delta < 1$ is the depreciation rate of capital and $\xi \geq 0$ is a parameter capturing adjustment costs related to physical capital.

For our numerical solutions, we use the usual functional form:

$$u_{i,t}\left(c_{i,t}, n_{i,t}, m_{i,t}, g_t\right) = \frac{c_{i,t}^{1-\sigma}}{1-\sigma} - \chi_n \frac{n_{i,t}^{1+\eta}}{1+\eta} + \chi_g \frac{g_t^{1-\zeta}}{1-\zeta}, \tag{10}$$

where χ_n, χ_g, σ, η, ζ are preference parameters.

Household's Optimality Conditions

Each household i acts competitively, taking prices and policy as given. Following the literature, to solve the household's problem, we follow a two-step procedure. We first suppose that the household determines its desired consumption of composite goods, $c_{i,t}^H$ and $c_{i,t}^F$, and, in turn, chooses how to distribute its purchases of individual varieties, $c_{i,t}^H(h)$ and $c_{i,t}^F(f)$.

The first-order conditions of each i include the budget constraints written above and also

$$\frac{\partial u_{i,t}}{\partial c_{i,t}}\frac{\partial c_{i,t}}{\partial c_{i,t}^H}\frac{P_t}{P_t^H\left(1+\tau_t^c\right)}$$
$$=\beta E_t\frac{\partial u_{i,t+1}}{\partial c_{i,t+1}}\frac{\partial c_{i,t+1}}{\partial c_{i,t+1}^H}\frac{P_{t+1}}{P_{t+1}^H\left(1+\tau_{t+1}^c\right)}R_t\frac{P_t}{P_{t+1}} \tag{11}$$

$$\frac{\partial u_{i,t}}{\partial c_{i,t}}\frac{\partial c_{i,t}}{\partial c_{i,t}^H}\frac{P_t}{P_t^H\left(1+\tau_t^c\right)}\frac{S_tP_t^*}{P_t}\left[1+\phi^h\left(\frac{S_tP_t^*}{P_t}f_t^h-\frac{SP^*}{P}f^h\right)\right]$$
$$=\beta E_t\frac{\partial u_{i,t+1}}{\partial c_{i,t+1}}\frac{\partial c_{i,t+1}}{\partial c_{t+1}^H}\frac{P_{t+1}}{P_{t+1}^H\left(1+\tau_{t+1}^c\right)}Q_t\frac{S_{t+1}P_{t+1}^*}{P_{t+1}}\frac{P_t^*}{P_{t+1}^*} \tag{12}$$

$$\frac{\partial u_{i,t}}{\partial c_{i,t}}\frac{\partial c_{i,t}}{\partial c_{i,t}^H}\frac{1}{\left(1+\tau_t^c\right)}\left\{1+\xi\left(\frac{k_{i,t}}{k_{i,t-1}}-1\right)\right\}$$
$$=\beta E_t\frac{\partial u_{i,t+1}}{\partial c_{i,t+1}}\frac{\partial c_{i,t+1}}{\partial c_{i,t+1}^H}\frac{1}{\left(1+\tau_{t+1}^c\right)}$$
$$\times\left\{\begin{array}{c}(1-\delta)-\frac{\xi}{2}\left(\frac{k_{i,t+1}}{k_{i,t}}-1\right)^2+\\ \xi\left(\frac{k_{i,t+1}}{k_{i,t}}-1\right)\frac{k_{i,t+1}}{k_{i,t}}+\left(1-\tau_{t+1}^k\right)r_{t+1}^k\end{array}\right\} \tag{13}$$

$$-\frac{\partial u_{i,t}}{\partial n_{i,t}}=\frac{\partial u_{i,t}}{\partial c_{i,t}}\frac{\partial c_{i,t}}{\partial c_{i,t}^H}\frac{P_t}{P_t^H}\frac{\left(1-\tau_t^n\right)w_t}{\left(1+\tau_t^c\right)} \tag{14}$$

$$\frac{c_{i,t}^H}{c_{i,t}^F}=\frac{\nu}{1-\nu}\frac{P_t^F}{P_t^H} \tag{15}$$

$$c_{i,t}^H(h)=\left[\frac{P_t^H(h)}{P_t^H}\right]^{-\phi}c_{i,t}^H \tag{16}$$

$$c_{i,t}^F(f)=\left[\frac{P_t^F(f)}{P_t^F}\right]^{-\phi}c_{i,t}^F. \tag{17}$$

Equations (11)–(13) are, respectively, the Euler equations for domestic bonds, foreign assets, and domestic capital, and (14) is the optimality condition for work hours. Finally, (15) shows the optimal allocation between domestic and foreign goods, while (16) and (17) show the optimal demand for each variety of domestic and foreign goods, respectively; these conditions are used in the firm's optimization problem below.

Implications for Price Bundles

Equations (15), (16), and (17), combined with the household's budget constraints, imply that the three price indexes are

$$P_t = (P_t^H)^\nu (P_t^F)^{1-\nu} \tag{18}$$

$$P_t^H = \left[\sum_{h=1}^{N} [P_t^H(h)]^{1-\phi}\right]^{\frac{1}{1-\phi}} \tag{19}$$

$$P_t^F = \left[\sum_{f=1}^{N} [P_t^F(f)]^{1-\phi}\right]^{\frac{1}{1-\phi}}. \tag{20}$$

Appendix 2. Firms

This appendix presents and solves the problem of the firm in some detail. There are $h = 1, 2, \ldots, N$ domestic firms. Each firm h produces a differentiated good of variety h under monopolistic competition in its own product market facing Calvo-type nominal price fixities.

Demand for the Firm's Product

Demand for firm h's product, $y_t^H(h)$, comes from domestic households' consumption and investment, $C_t^H(h)$ and $X_t(h)$, where $C_t^H(h) \equiv \sum_{i=1}^{N} c_{i,t}^H(h)$ and $X_t(h) \equiv \sum_{i=1}^{N} x_{i,t}(h)$, from the domestic government, $G_t(h)$, and from foreign households' consumption, $C_t^{F*}(h) \equiv \sum_{i=1}^{N*} c_{i,t}^{F*}(h)$. Thus, the demand for each domestic firm's product is

$$y_t^H(h) = C_t^H(h) + X_t(h) + G_t(h) + C_t^{F*}(h). \tag{21}$$

Using demand functions like those derived by solving the household's problem above, we have

$$c_{i,t}^H(h) = \left[\frac{P_t^H(h)}{P_t^H}\right]^{-\phi} c_{i,t}^H \tag{22}$$

$$x_{i,t}(h) = \left[\frac{P_t^H(h)}{P_t^H}\right]^{-\phi} x_{i,t} \tag{23}$$

$$G_t(h) = \left[\frac{P_t^H(h)}{P_t^H}\right]^{-\phi} G_t \tag{24}$$

$$c_{i,t}^{F*}(h) = \left[\frac{P_t^{F*}(h)}{P_t^{F*}}\right]^{-\phi} c_{i,t}^{F*}, \tag{25}$$

where, using the law of one price, we have in (25)

$$\frac{P_t^{F*}(h)}{P_t^{F*}} = \frac{\frac{P_t^H(h)}{S_t}}{\frac{P_t^H}{S_t}} = \frac{P_t^H(h)}{P_t^H}. \tag{26}$$

Since, at the economy level, aggregate demand is

$$Y_t^H = C_t^H + X_t + G_t + C_t^{F*}, \tag{27}$$

the above equations imply that the demand for each firm's product is

$$y_t^H(h) = \left[\frac{P_t^H(h)}{P_t^H}\right]^{-\phi} Y_t^H. \tag{28}$$

The Firm's Problem

The real profit of each firm h is (see, e.g., Benigno and Thoenissen 2008 for similar modeling)

$$\widetilde{\omega}_t(h) = \frac{P_t^H(h)}{P_t} y_t^H(h) - \frac{P_t^H}{P_t} r_t^k k_{t-1}(h) - w_t n_t(h). \tag{29}$$

The production technology is

$$y_t^H(h) = A_t[k_{t-1}(h)]^{\alpha}[n_t(h)]^{1-\alpha}, \tag{30}$$

where A_t is an exogenous stochastic total factor productivity process whose motion is defined below.

Profit maximization is also subject to the demand for its product as derived above:

$$y_t^H(h) = \left[\frac{P_t^H(h)}{P_t^H}\right]^{-\phi} Y_t^H. \tag{31}$$

In addition, firms choose their prices facing a Calvo-type nominal fixity. In each period, firm h faces an exogenous probability θ of not being able to reset its price. A firm h, which is able to reset its price, chooses its price $P_t^{\#}(h)$ to maximize the sum of discounted expected nominal profits for the next k periods in which it may have to keep its price fixed. This is modeled next.

Firm's Optimality Conditions

To solve the firm's problem, we follow a two-step procedure. We first solve a cost-minimization problem, where each firm h minimizes its cost by choosing factor inputs given technology and prices. The solution will give a minimum nominal cost function, which is a function of factor prices and output produced by the firm. In turn, given this cost function, each firm, which is able to reset its price, solves a maximization problem by choosing its price.

The solution to the cost-minimization problem gives the input demand functions:

$$w_t = mc_t(h)(1-a)\frac{y_t(h)}{n_t(h)} \tag{32}$$

$$\frac{P_t^H}{P_t} r_t^k = mc_t(h) a \frac{y_t(h)}{k_{t-1}(h)}, \tag{33}$$

where $mc_t(h) = \frac{\Psi_t'(h)}{P_t}$ denotes the real marginal cost.

Then, the firm chooses its price to maximize nominal profits written as (see also, e.g., Rabanal 2009 for similar modeling)

$$E_t \sum_{k=0}^{\infty} \theta^k \Xi_{t,t+k} \left\{ P_t^{\#}(h)\, y_{t+k}^H(h) - \Psi_{t+k}\left(y_{t+k}^H(h)\right) \right\},$$

where $\Xi_{t,t+k}$ is a discount factor taken as given by the firm, $y_{t+k}^H(h) = \left[\frac{P_t^{\#}(h)}{P_{t+k}^H}\right]^{-\phi} Y_{t+k}^H$ and $\Psi_t(h)$ denotes the minimum nominal cost function for producing $y_t^H(h)$ at t so that $\Psi_t'(h)$ is the associated marginal cost.

The first-order condition gives

$$E_t \sum_{k=0}^{\infty} \theta^k \Xi_{t,t+k} \left[\frac{P_t^{\#}(h)}{P_{t+k}^H} \right]^{-\phi} Y_{t+k}^H \left\{ P_t^{\#}(h) - \frac{\phi}{\phi - 1} \Psi'_{t+k}(h) \right\} = 0. \tag{34}$$

Dividing by P_t^H, we have

$$E_t \sum_{k=0}^{\infty} \theta^k \left[\Xi_{t,t+k} \left[\frac{P_t^{\#}(h)}{P_{t+k}^H} \right]^{-\phi} Y_{t+k}^H \left\{ \frac{P_t^{\#}(h)}{P_t^H} - \frac{\phi}{\phi - 1} mc_{t+k}(h) \frac{P_{t+k}}{P_t^H} \right\} \right] = 0. \tag{35}$$

Summing up, the behavior of each firm h is summarized by (32), (33), and (35). A recursive expression of the equation above is presented below.

Notice, finally, that each firm h, which can reset its price in period t, solves an identical problem, so $P_t^{\#}(h) = P_t^{\#}$ is independent of h, and each firm h, which cannot reset its price, just sets its previous period price $P_t^H(h) = P_{t-1}^H(h)$. Thus, the evolution of the aggregate price level is given by

$$\left(P_t^H\right)^{1-\phi} = \theta \left(P_{t-1}^H\right)^{1-\phi} + (1-\theta) \left(P_t^{\#}\right)^{1-\phi}. \tag{36}$$

Appendix 3. Government Budget Constraint

This appendix presents in some detail the government budget constraint and the menu of fiscal policy instruments. Recall that we have assumed a cashless economy for simplicity.

Then, the period budget constraint of the government written in aggregate and nominal quantities is

$$\begin{aligned} B_t + S_t F_t^g = {} & P_t \frac{\phi^g}{2} \left(\frac{S_t F_t^g}{P_t} - \frac{S F^g}{P} \right)^2 + R_{t-1} B_{t-1} + Q_{t-1} S_t F_{t-1}^g \\ & + P_t^H G_t - \tau_t^c (P_t^H C_t^H + P_t^F C_t^F) \\ & - \tau_t^k (r_t^k P_t^H K_{t-1} + P_t \widetilde{\Omega}_t) - \tau_t^n W_t \widetilde{N}_t - T_t^l, \end{aligned} \tag{37}$$

where B_t is the end-of-period nominal domestic public debt and F_t^g is the end-of-period nominal foreign public debt expressed in foreign currency so it is multiplied by the exchange rate, S_t. The parameter $\phi^g \geq 0$ captures adjustment costs related to public foreign debt. The rest of the notation is as above. Note that $B_t = \sum_{i=1}^N B_{i,t}$, $C_t^H = \sum_{i=1}^N c_{i,t}^H$, $C_t^F = \sum_{i=1}^N c_{i,t}^F$, $K_{t-1} = \sum_{i=1}^N k_{i,t-1}$, $\widetilde{\Omega}_t \equiv \sum_{i=1}^N \widetilde{\omega}_{i,t}$, $\widetilde{N}_t \equiv \sum_{i=1}^N n_{i,t}$, and T_t^l denotes the nomimal value of lump-sum taxes.

Let $D_t = B_t + S_t F_t^g$ denote the total nominal public debt at the end of the period. This can be held both by domestic private agents, $\lambda_t D_t$, where in equilibrium $B_t \equiv \lambda_t D_t$, and by foreign private agents, $S_t F_t^g \equiv (1-\lambda_t) D_t$, where $0 \leq \lambda_t \leq 1$. Then, dividing by the price level and the number of households, the budget constraint in real and per capita terms is

$$\begin{aligned} d_t = {} & \frac{\phi^g}{2}[(1-\lambda_t) d_t - (1-\lambda) d]^2 + R_{t-1}\frac{P_{t-1}}{P_t}\lambda_{t-1}d_{t-1} \\ & + Q_{t-1}\frac{S_t P_t^*}{P_t}\frac{P_{t-1}^*}{P_t^*}\frac{P_{t-1}}{P_{t-1}^* S_{t-1}}(1-\lambda_{t-1}) d_{t-1} \\ & + \frac{P_t^H}{P_t} g_t - \tau_t^c \left(\frac{P_t^H}{P_t} c_t^H + \frac{P_t^F}{P_t} c_t^F\right) \\ & - \tau_t^k \left(r_t^k \frac{P_t^H}{P_t} k_{t-1} + \widetilde{\omega}_t\right) - \tau_t^n w_t n_t - \tau_t^l, \end{aligned} \tag{38}$$

where, as said in the main text, in each period, one of the policy instruments (τ_t^c, τ_t^k, τ_t^n, g_t, τ_t^l, λ_t, d_t) follows residually to satisfy the government budget constraint.

Appendix 4. Decentralized Equilibrium (DE)

This appendix presents in some detail the DE system. Following the related literature, we work in steps.

Equilibrium Equations

The DE is summarized by the following equations (quantities are in per capita terms):

$$\frac{\partial u_t}{\partial c_t}\frac{\partial c_t}{\partial c_t^H}\frac{1}{(1+\tau_t^c)}\frac{P_t}{P_t^H}=\beta E_t\frac{\partial u_{t+1}}{\partial c_{t+1}}\frac{\partial c_{t+1}}{\partial c_{t+1}^H}\frac{1}{\left(1+\tau_{t+1}^c\right)}\frac{P_{t+1}}{P_{t+1}^H}R_t\frac{P_t}{P_{t+1}} \tag{39}$$

$$\begin{aligned}&\frac{\partial u_t}{\partial c_t}\frac{\partial c_t}{\partial c_t^H}\frac{1}{(1+\tau_t^c)}\frac{P_t}{P_t^H}\frac{S_tP_t^*}{P_t}\left(1+\phi^p\left(\frac{S_tP_t^*}{P_t}f_t^h-\frac{SP^*}{P}f^h\right)\right)\\&\quad=\beta\frac{\partial u_{t+1}}{\partial c_{t+1}}\frac{\partial c_{t+1}}{\partial c_{t+1}^H}\frac{1}{\left(1+\tau_{t+1}^c\right)}\frac{P_{t+1}}{P_{t+1}^H}Q_t\frac{S_{t+1}P_{t+1}^*}{P_{t+1}}\frac{P_t^*}{P_{t+1}^*}\end{aligned} \tag{40}$$

$$\begin{aligned}&\frac{\partial u_t}{\partial c_t}\frac{\partial c_t}{\partial c_t^H}\frac{1}{(1+\tau_t^c)}\left\{1+\xi\left(\frac{k_t}{k_{t-1}}-1\right)\right\}\\&\quad=\beta E_t\frac{\partial u_{t+1}}{\partial c_{t+1}}\frac{\partial c_{t+1}}{\partial c_{t+1}^H}\frac{1}{\left(1+\tau_{t+1}^c\right)}\\&\qquad\times\left\{\begin{array}{c}(1-\delta)-\frac{\xi}{2}\left(\frac{k_{t+1}}{k_t}-1\right)^2\\+\xi\left(\frac{k_{t+1}}{k_t}-1\right)\frac{k_{t+1}}{k_t}+\left(1-\tau_{t+1}^k\right)r_{t+1}^k\end{array}\right\}\end{aligned} \tag{41}$$

$$-\frac{\partial u_t}{\partial n_t}=\frac{\partial u_t}{\partial c_t}\frac{\partial c_t}{\partial c_t^H}\frac{P_t}{P_t^H}\frac{(1-\tau_t^n)\,w_t}{(1+\tau_t^c)} \tag{42}$$

$$\frac{c_t^H}{c_t^F}=\frac{\nu}{1-\nu}\frac{P_t^F}{P_t^H} \tag{43}$$

$$k_t=(1-\delta)k_{t-1}+x_t-\frac{\xi}{2}\left(\frac{k_t}{k_{t-1}}-1\right)^2k_{t-1} \tag{44}$$

$$c_t\equiv\frac{\left(c_t^H\right)^\nu\left(c_t^F\right)^{1-\nu}}{\nu^\nu\left(1-\nu\right)^{1-\nu}} \tag{45}$$

$$w_t=mc_t(1-a)A_tk_{t-1}^an_t^{-a} \tag{46}$$

$$\frac{P_t^H}{P_t}r_t^k=mc_taA_tk_{t-1}^{a-1}n_t^{1-a} \tag{47}$$

$$\widetilde{\omega}_t=\frac{P_t^H}{P_t}y_t^H-\frac{P_t^H}{P_t}r_t^kk_{t-1}-w_tn_t \tag{48}$$

$$E_t\sum_{k=0}^{\infty}\theta^k\left[\Xi_{t,t+k}\left[\frac{P_t^{\#}}{P_{t+k}^H}\right]^{-\phi}y_{t+k}^H\left\{\frac{P_t^{\#}}{P_t}-\frac{\phi}{\phi-1}mc_{t+k}\frac{P_{t+k}}{P_t}\right\}\right]=0 \tag{49}$$

$$y_t^H = \frac{1}{\left[\frac{\widetilde{P}_t^H}{P_t^H}\right]^{-\phi}} A_t k_{t-1}^a n_t^{1-a} \tag{50}$$

$$\begin{aligned} d_t = {} & R_{t-1}\frac{P_{t-1}}{P_t}\lambda_{t-1}d_{t-1} + Q_{t-1}\frac{S_t P_t^*}{P_t}\frac{P_{t-1}^*}{P_t^*}\frac{P_{t-1}}{P_{t-1}^* S_{t-1}}\left(1-\lambda_{t-1}\right)d_{t-1} \\ & + \frac{P_t^H}{P_t}g_t - \tau_t^c\left(\frac{P_t^H}{P_t}c_t^H + \frac{P_t^F}{P_t}c_t^F\right) - \tau_t^k\left(r_t^k\frac{P_t^H}{P_t}k_{t-1} + \widetilde{\omega}_t\right) \\ & - \tau_t^n w_t n_t - \tau_t^l + \frac{\phi^g}{2}\left[\left(1-\lambda_t\right)d_t - \left(1-\lambda\right)d\right]^2 \end{aligned} \tag{51}$$

$$y_t^H = c_t^H + x_t + g_t + c_t^{F*} \tag{52}$$

$$\begin{aligned} & -\frac{P_t^H}{P_t}c_t^{F*} + \frac{P_t^F}{P_t}c_t^F + \frac{\phi^g}{2}\left[\left(1-\lambda_t\right)d_t - \left(1-\lambda\right)d\right]^2 \\ & + \frac{\phi^p}{2}\left(\frac{S_t P_t^*}{P_t}f_t^h - \frac{SP^*}{P}f^h\right)^2 Q_{t-1}\frac{S_t P_t^*}{P_t}\frac{P_{t-1}^*}{P_t^*} \\ & \times\left(\frac{\left(1-\lambda_{t-1}\right)d_{t-1}}{\frac{S_{t-1}P_{t-1}^*}{P_{t-1}}} - f_{t-1}^h\right) = \left(1-\lambda_t\right)d_t - \frac{S_t P_t^*}{P_t}f_t^h \end{aligned} \tag{53}$$

$$\left(P_t^H\right)^{1-\phi} = \theta P_{t-1}^{1-\phi} + \left(1-\theta\right)\left(P_t^{\#}\right)^{1-\phi} \tag{54}$$

$$P_t = \left(P_t^H\right)^{\nu}\left(P_t^F\right)^{1-\nu} \tag{55}$$

$$P_t^F = S_t P_t^{H*} \tag{56}$$

$$P_t^* = \left(P_t^{H*}\right)^{\nu^*}\left(\frac{P_t^H}{S_t}\right)^{1-\nu^*} \tag{57}$$

$$\left(\widetilde{P}_t^H\right)^{-\phi} = \theta\left(\widetilde{P}_{t-1}^H\right)^{-\phi} + \left(1-\theta\right)\left(P_t^{\#}\right)^{-\phi} \tag{58}$$

$$Q_t = Q_t^* + \psi\left(e^{\frac{d_t}{\frac{P_t^H}{P_t}Y_t^H} - \overline{d}} - 1\right) \tag{59}$$

$$l_t \equiv \frac{R_t\lambda_t d_t + Q_t\frac{S_{t+1}}{S_t}(1-\lambda_t)d_t}{\frac{P_t^H}{P_t}y_t^H}, \tag{60}$$

where $f_t^{*g} \equiv \frac{(1-\lambda_t)d_t}{\frac{S_t P_t^*}{P_t}}$, $\Xi_{t,t+k} \equiv \beta^k \frac{c_{t+k}^{-\sigma}}{c_t^{-\sigma}} \frac{P_t}{P_{t+k}} \frac{\tau_t^c}{\tau_{t+k}^c}$, $y_t^H = \left[\sum_{h=1}^{N} [y_t^H(h)]^{\frac{\phi-1}{\phi}}\right]^{\frac{\phi}{\phi-1}}$, and $\widetilde{P}_t^H \equiv \left(\sum_{h=1}^{N} [P_t(h)]^{-\phi}\right)^{-\frac{1}{\phi}}$. Thus, $\left(\frac{\widetilde{P}_t^H}{P_t^H}\right)^{-\phi}$ is a measure of price dispersion.

We thus have twenty-two equations in twenty-two variables, $\{y_t^H$, c_t, c_t^H, c_t^F, n_t, x_t, k_t, f_t^h, P_t^F, P_t, P_t^H, $P_t^{\#}$, $\widetilde{P}_t^H$, w_t, mc_t, $\widetilde{\omega}_t$, r_t^k, Q_t, d_t, P_t^*, R_t, $l_t\}_{t=0}^{\infty}$. This is given the independently set monetary and fiscal policy instruments, $\{S_t,\ \tau_t^c,\ \tau_t^k,\ \tau_t^n,\ g_t,\ \tau_t^l,\ \lambda_t\}_{t=0}^{\infty}$; the rest-of-the-world variables, $\{c_t^{F*}, Q_t^*, P_t^{H*}\}_{t=0}^{\infty}$; technology, $\{A_t\}_{t=0}^{\infty}$; and initial conditions for the state variables.

In what follows, we shall transform the above equilibrium conditions. In particular, following the related literature, we rewrite them—first, by expressing price levels in inflation rates, secondly, by writing the firm's optimality conditions in recursive form, and, thirdly, by introducing a new equation that helps us to compute expected discounted lifetime utility. Finally, we will present the final transformed system that is solved numerically.

Variables Expressed in Ratios

We first express prices in rate form. We define seven new variables, which are the gross domestic CPI inflation rate, $\Pi_t \equiv \frac{P_t}{P_{t-1}}$; the gross foreign CPI inflation rate, $\Pi_t^* \equiv \frac{P_t^*}{P_{t-1}^*}$; the gross domestic goods inflation rate, $\Pi_t^H \equiv \frac{P_t^H}{P_{t-1}^H}$; the auxiliary variable, $\Theta_t \equiv \frac{P_t^{\#}}{P_t^H}$; the price dispersion index, $\Delta_t \equiv \left[\frac{\widetilde{P}_t^H}{P_t^H}\right]^{-\phi}$; the gross rate of exchange rate depreciation, $\epsilon_t \equiv \frac{S_t}{S_{t-1}}$; and the terms of trade, $TT_t \equiv \frac{P_t^F}{P_t^H} = \frac{S_t P_t^{*H}}{P_t^H}$.[1] In what follows, we use $\Pi_t, \Pi_t^*, \Pi_t^H, \Theta_t, \Delta_t, \epsilon_t,\ TT_t$ instead of $P_t, P_t^*, P_t^H, P_t^{\#}, \widetilde{P}_t, S_t,\ P_t^F$, respectively.

Also, for convenience and comparison with the data, we express fiscal and public finance variables as shares of nominal output,

[1]Thus, $\frac{TT_t}{TT_{t-1}} = \frac{\frac{S_t}{S_{t-1}} \frac{P_t^{*H}}{P_{t-1}^{*H}}}{\frac{P_t^H}{P_{t-1}^H}} = \frac{\epsilon_t \Pi_t^{*H}}{\Pi_t^H}$.

$P_t^H y_t^H$. In particular, using the definitions above, real government spending, g_t, can be written as $g_t = s_t^g y_t^H$, while real government transfers, τ_t^l, can be written as $\tau_t^l = s_t^l y_t^H TT_t^{\nu-1}$, where s_t^g and s_t^l denote, respectively, the output shares of government spending and government transfers.

Equation (49) Expressed in Recursive Form

We now replace equation (49), from the firm's optimization problem, with an equivalent equation in recursive form. In particular, following Schmitt-Grohé and Uribe (2007), we look for a recursive representation of the form

$$E_t \sum_{k=0}^{\infty} \theta^k \Xi_{t,t+k} \left[\frac{P_t^{\#}}{P_{t+k}^H} \right]^{-\phi} y_{t+k}^H \left\{ P_t^{\#} - \frac{\phi}{(\phi-1)} mc_{t+k} P_{t+k} \right\} = 0. \tag{61}$$

We define two auxiliary endogenous variables:

$$z_t^1 \equiv E_t \sum_{k=0}^{\infty} \theta^k \Xi_{t,t+k} \left[\frac{P_t^{\#}}{P_{t+k}^H} \right]^{-\phi} y_{t+k}^H \frac{P_t^{\#}}{P_t} \tag{62}$$

$$z_t^2 \equiv E_t \sum_{k=0}^{\infty} \theta^k \Xi_{t,t+k} \left[\frac{P_t^{\#}}{P_{t+k}^H} \right]^{-\phi} y_{t+k}^H mc_{t+k} \frac{P_{t+k}}{P_t}. \tag{63}$$

Using these two auxiliary variables, z_t^1 and z_t^2, as well as equation (61), we come up with two new equations that enter the dynamic system and allow a recursive representation of (61).

Thus, in what follows, we replace equation (49) with

$$z_t^1 = \frac{\phi}{(\phi-1)} z_t^2, \tag{64}$$

where we also have the two new equations (and the two new endogenous variables, z_t^1 and z_t^2)

$$z_t^1 = \Theta_t^{1-\phi} y_t TT_t^{\nu-1} + \beta\theta E_t \frac{c_{t+1}^{-\sigma}}{c_t^{-\sigma}} \frac{1+\tau_t^c}{1+\tau_{t+1}^c} \left(\frac{\Theta_t}{\Theta_{t+1}} \right)^{1-\phi} \left(\frac{1}{\Pi_{t+1}^H} \right)^{1-\phi} z_{t+1}^1 \tag{65}$$

$$z_t^2 = \Theta_t^{-\phi} y_t mc_t + \beta\theta E_t \frac{c_{t+1}^{-\sigma}}{c_t^{-\sigma}} \frac{1+\tau_t^c}{1+\tau_{t+1}^c} \left(\frac{\Theta_t}{\Theta_{t+1}}\right)^{-\phi} \left(\frac{1}{\Pi_{t+1}^H}\right)^{-\phi} z_{t+1}^2. \tag{66}$$

Lifetime Utility Written as a First-Order Difference Equation

Since we want to compute social welfare, we follow Schmitt-Grohé and Uribe (2007) by defining a new endogenous variable, V_t, whose motion is given by

$$V_t = \frac{c_t^{1-\sigma}}{1-\sigma} - \chi_n \frac{n_t^{1+\eta}}{1+\eta} + \chi_g \frac{\left(s_t^g y_t^H\right)^{1-\zeta}}{1-\zeta} + \beta E_t V_{t+1}, \tag{67}$$

where V_t is household's expected discounted lifetime utility at time t.

Thus, in what follows, we add equation (67) and the new variable V_t to the equilibrium system. Note that, in the welfare numerical solutions reported, we add a constant number, 2, to each period's utility; this makes the welfare numbers easier to read.

Equilibrium Equations Transformed

Using the above, the transformed DE system is summarized by

$$V_t = \frac{c_t^{1-\sigma}}{1-\sigma} - \chi_n \frac{n_t^{1+\eta}}{1+\eta} + \chi_g \frac{\left(s_t^g y_t^H\right)^{1-\zeta}}{1-\zeta} + \beta E_t V_{t+1} \tag{68}$$

$$\beta E_t \; c_{t+1}^{-\sigma} \frac{1}{\left(1+\tau_{t+1}^c\right)} R_t \frac{1}{\Pi_{t+1}} = c_t^{-\sigma} \frac{1}{\left(1+\tau_t^c\right)} \tag{69}$$

$$\begin{aligned} &\beta E_t c_{t+1}^{-\sigma} \frac{1}{\left(1+\tau_{t+1}^c\right)} Q_t TT_{t+1}^{v^*+\nu-1} \frac{1}{\Pi_{t+1}^*} \\ &\quad = c_t^{-\sigma} \frac{1}{\left(1+\tau_t^c\right)} TT_t^{v^*+\nu-1} \left[1 + \phi^p \left(TT_t^{\nu^*+\nu-1} f_t^h - TT^{\nu^*+\nu-1} f^h\right)\right] \end{aligned} \tag{70}$$

$$\begin{aligned} &\beta c_{t+1}^{-\sigma} TT_{t+1}^{\nu-1} \frac{1}{\left(1+\tau_{t+1}^c\right)} \left\{1 - \delta - \frac{\xi}{2}\left(\frac{k_{t+1}}{k_t} - 1\right)^2 \right. \\ &\quad \left. + \xi \left(\frac{k_{t+1}}{k_t} - 1\right) \frac{k_{t+1}}{k_t} + \left(1 - \tau_{t+1}^k\right) r_{t+1}^k \right\} \\ &\quad = c_t^{-\sigma} TT_t^{\nu-1} \frac{1}{\left(1+\tau_t^c\right)} \left[1 + \xi \left(\frac{k_t}{k_{t-1}} - 1\right)\right] \end{aligned} \tag{71}$$

$$\chi_n n_t^{\eta} = c_t^{-\sigma} \frac{(1-\tau_t^n)\, w_t}{(1+\tau_t^c)} \tag{72}$$

$$\frac{c_t^H}{c_t^F} = \frac{\nu}{1-\nu} TT_t \tag{73}$$

$$k_t = (1-\delta)k_{t-1} + x_t - \frac{\xi}{2}\left(\frac{k_t}{k_{t-1}} - 1\right)^2 k_{t-1} \tag{74}$$

$$c_t \equiv \frac{\left(c_t^H\right)^{\nu}\left(c_t^F\right)^{1-\nu}}{(\nu)^{\nu}(1-\nu)^{1-\nu}} \tag{75}$$

$$w_t = mc_t(1-a)A_t k_{t-1}^a n_t^{-a} \tag{76}$$

$$\frac{1}{TT_t^{1-v}} r_t^k = mc_t a A_t k_{t-1}^{a-1} n_t^{1-a} \tag{77}$$

$$\widetilde{\omega}_t = \frac{1}{TT_t^{1-v}} y_t^H - \frac{1}{TT_t^{1-v}} r_t^k k_{t-1} - w_t n_t \tag{78}$$

$$z_t^1 = \frac{\phi}{(\phi-1)} z_t^2 \tag{79}$$

$$y_t^H = \frac{1}{\Delta_t} A_t k_{t-1}^a n_t^{1-a} \tag{80}$$

$$\begin{aligned} d_t = {} & \frac{\phi^g}{2}[(1-\lambda_t)d_t - (1-\lambda)d]^2 + R_{t-1}\frac{1}{\Pi_t}\lambda_{t-1}d_{t-1} \\ & + Q_{t-1}TT_t^{v+v^*-1}\frac{1}{\Pi_t^*}\frac{1}{TT_{t-1}^{v+v^*-1}}(1-\lambda_{t-1})d_{t-1} + TT_t^{\nu-1}s_t^g y_t^H \\ & - \tau_t^c\left(\frac{1}{TT_t^{1-v}}c_t^H + TT_t^v c_t^F\right) - \tau_t^k\left(r_t^k\frac{1}{TT_t^{1-v}}k_{t-1} + \widetilde{\omega}_t\right) \\ & - \tau_t^n w_t n_t - TT_t^{\nu-1}s_t^l y_t^H \end{aligned} \tag{81}$$

$$y_t^H = c_t^H + x_t + s_t^g y_t^H + c_t^{F*} \tag{82}$$

$$\begin{aligned} (1-\lambda_t)d_t - TT_t^{\nu^*+\nu-1}f_t^h = {} & -TT_t^{\nu-1}c_t^{F*} + TT_t^{\nu}c_t^F \\ & + Q_{t-1}TT_t^{\nu^*+\nu-1}\frac{1}{\Pi_t^*}\left(\frac{1}{TT_{t-1}^{v+v^*-1}}(1-\lambda_{t-1})d_{t-1} - f_{t-1}^h\right) \end{aligned}$$

$$+\frac{\phi^p}{2}\left(TT_t^{\nu^*+\nu-1}f_t^h - TT^{\nu^*+\nu-1}f^h\right)^2$$
$$+\frac{\phi^g}{2}\left((1-\lambda_t)d_t-(1-\lambda)d\right)^2 \tag{83}$$
$$\left(\Pi_t^H\right)^{1-\phi} = \theta + (1-\theta)\left(\Theta_t\Pi_t^H\right)^{1-\phi} \tag{84}$$
$$\frac{\Pi_t}{\Pi_t^H} = \left(\frac{TT_t}{TT_{t-1}}\right)^{1-\nu} \tag{85}$$
$$\frac{TT_t}{TT_{t-1}} = \frac{\epsilon_t \Pi_t^{H*}}{\Pi_t^H} \tag{86}$$
$$\frac{\Pi_t^*}{\Pi_t^{H*}} = \left(\frac{TT_{t-1}}{TT_t}\right)^{1-\nu^*} \tag{87}$$
$$\Delta_t = \theta\Delta_{t-1}\left(\Pi_t^H\right)^{\phi} + (1-\theta)\left(\Theta_t\right)^{-\phi} \tag{88}$$
$$Q_t = Q_t^* + \psi\left(e^{\left(\frac{d_t}{TT_t^{\nu-1}y_t^H}-\bar{d}\right)} - 1\right) \tag{89}$$
$$z_t^1 = \Theta_t^{1-\phi}y_t TT_t^{\nu-1} + \beta\theta E_t\ \frac{c_{t+1}^{-\sigma}}{c_t^{-\sigma}}\frac{1+\tau_t^c}{1+\tau_{t+1}^c}\left(\frac{\Theta_t}{\Theta_{t+1}}\right)^{1-\phi}\left(\frac{1}{\Pi_{t+1}^H}\right)^{1-\phi} z_{t+1}^1 \tag{90}$$
$$z_t^2 = \Theta_t^{-\phi}y_t mc_t + \beta\theta E_t\ \frac{c_{t+1}^{-\sigma}}{c_t^{-\sigma}}\frac{1+\tau_t^c}{1+\tau_{t+1}^c}\left(\frac{\Theta_t}{\Theta_{t+1}}\right)^{-\phi}\left(\frac{1}{\Pi_{t+1}^H}\right)^{-\phi} z_{t+1}^2 \tag{91}$$
$$l_t = \frac{R_t s_t d_t + Q_t\epsilon_{t+1}\left(1-\lambda_t\right)d_t}{TT_t^{\nu-1}y_t^H}. \tag{92}$$

We thus have twenty-five equations in twenty-five variables, $\{V_t, y_t^H, c_t,\ c_t^H, c_t^F, n_t, x_t, k_t, f_t^h,\ TT_t, \Pi_t, \Pi_t^H, \Theta_t, \Delta_t, w_t, mc_t, \widetilde{\omega}_t, r_t^k, Q_t, d_t, \Pi_t^*, z_t^1, z_t^2, R_t,\ l_t\}_{t=0}^{\infty}$. This is given the independently set policy instruments, $\{\epsilon_t, \tau_t^c,\ \tau_t^k, \tau_t^n, s_t^g, s_t^l,\ \lambda_t\}_{t=0}^{\infty}$; the rest-of-the-world variables, $\{c_t^{F*}, Q_t^*, \Pi_t^{H*}\}_{t=0}^{\infty}$; technology, $\{A_t\}_{t=0}^{\infty}$; and initial conditions for the state variables.

Equations and Unknowns (Given Feedback Policy Coefficients)

We can now define the final equilibrium system. It consists of the twenty-five equations of the transformed DE presented in the previous subsection, the four policy rules in subsection 2.7 in the main text, and the equation for domestic exports in subsection 2.8 in the main text. Thus, we have a system of thirty equations. By using two auxilliary variables, we transform it to a first order,[2] so we end up with thirty-two equations in thirty-two variables, $\{y_t^H$, c_t, c_t^H, c_t^F, x_t, n_t, f_t^h, d_t, k_t, $\widetilde{\omega}_t$, mc_t, Π_t, Π_t^H, Π_t^*, Θ_t, Δ_t, TT_t, w_t, r_t^k, Q_t, l_t, z_t^1, z_t^2, V_t, R_t; τ_t^c, s_t^g, τ_t^k, τ_t^n; $klead_t$, $TTlag_t$, $c_t^{F*}\}_{t=0}^{\infty}$. This is given the exogenous variables, $\{Q_t^*, \Pi_t^{H*}, A_t, \epsilon_t,\ s_t^l,\ \lambda_t\}_{t=0}^{\infty}$, and initial conditions for the state variables. The thirty-two endogenous variables are distinguished in twenty-four control variables, $\{y_t^H, c_t, c_t^H, c_t^F, x_t, n_t, \widetilde{\omega}_t, mc_t, \Pi_t, \Pi_t^H, \Pi_t^*, \Theta_t, TT_t, w_t, r_t^k, z_t^1, z_t^2, V_t, klead_t, c_t^{F*}, \tau_t^c,\ s_t^g,\ \tau_t^k,\ \tau_t^n\}_{t=0}^{\infty}$, and eight state variables, $\{f_{t-1}^h, d_{t-1}, k_{t-1}, \Delta_{t-1}, Q_{t-1}, R_{t-1}, l_{t-1}, TTlag_t\}_{t=0}^{\infty}$. All this is given the values of feedback policy coefficients in the policy rules, which are chosen optimally in the computational part of the paper.

Status Quo Steady State

In the steady state of the above model economy, if $\epsilon_t = 1$, our equations imply $\Pi_t^H = \Pi_t^{H*} = \Pi^* = \Pi$. If we also set $\Pi^{H*} = 1$ (this is the exogenous component of foreign inflation), then $\Pi_t^H = \Pi_t^{H*} = \Pi^* = \Pi = 1$; this in turn implies $\Delta = \Theta = 1$. Also, in order to solve the model numerically, we need a value for the exogenous exports, c^{F*}; we assume $c^{F*} = 1.01c^F$, as it is the case in the data. Finally, since from the Euler for bonds, $1 = \frac{\beta R}{\Pi}$, we residually get $R = \frac{1}{\beta}$. (Note that these conditions will also hold in all steady-state solutions throughout the paper.)

The steady-state system is (in the labor supply condition, we use $\frac{\nu c}{c^H} TT^{1-v} = 1$)

$$1 = \beta[1 - \delta + \left(1 - \tau^k\right) r^k] \tag{93}$$

$$1 = \beta Q \tag{94}$$

$$\chi_n n^\eta = c^{-\sigma} \frac{\nu c}{c^H} TT^{1-v} \frac{(1 - \tau^n)}{(1 + \tau^c)} w \tag{95}$$

[2] In particular, when we use the Schmitt-Grohé and Uribe (2004) MATLAB routines, we add two auxiliary endogenous variables, *klead* and *TTlag*, to reduce the dynamic system into a first-order one.

$$\frac{c^H}{c^F} = \frac{\nu}{1-\nu}TT \tag{96}$$

$$x = \delta k \tag{97}$$

$$c = \frac{\left(c^H\right)^\nu \left(c^F\right)^{1-\nu}}{(\nu)^\nu (1-\nu)^{1-\nu}} \tag{98}$$

$$w = mc(1-a)Ak^a n^{-a} \tag{99}$$

$$r^k = TT^{1-v} mcaAk^{a-1} n^{1-a} \tag{100}$$

$$\widetilde{\omega} = \frac{1}{TT^{1-v}} y^H - \frac{1}{TT^{1-v}} r^k k - wn \tag{101}$$

$$y^H = Ak^a n^{1-a} \tag{102}$$

$$\begin{aligned} d = Qd + TT^{\nu-1} s^g y^H - \tau^c \left(\frac{1}{TT^{1-v}} c^H + TT^v c^F \right) \\ - \tau^k \left(r^k \frac{1}{TT^{1-v}} k + \widetilde{\omega} \right) - \tau^n wn - TT^{\nu-1} s^l y^H \end{aligned} \tag{103}$$

$$y^H = c^H + x + s^g y^H + c^{F*} \tag{104}$$

$$\begin{aligned} (1-\lambda)d - TT^{\nu^*+\nu-1} f^h = -TT^{\nu-1} c^{F*} + TT^\nu c^F \\ + QTT^{\nu^*+\nu-1} \left(\frac{1}{TT^{v+v^*-1}} (1-\lambda)d - f^h \right) \end{aligned} \tag{105}$$

$$Q = Q^* + \psi \left(e^{\left(\frac{d}{TT^{\nu-1} y^H} - \overline{d} \right)} - 1 \right) \tag{106}$$

$$z^1 = \frac{\phi}{(\phi - 1)} z^2 \tag{107}$$

$$z^1 = TT^{\nu-1} y^H + \beta\theta z^1 \tag{108}$$

$$z^2 = y^H mc + \beta\theta z^2. \tag{109}$$

We thus have seventeen equations in seventeen variables, c, c^H, c^F, k, x, n, y_t^H, TT, $mc, \widetilde{\omega}, d, z^1, z^2$, Q, f^h, r^k, w. The numerical solution of this system is in table 2 in the main text.

Appendix 5. The Reformed Economy

This appendix presents the reformed economy as defined in subsection 4.1 in the main text. The equations describing the reformed economy are as in the previous appendix except that now we set $Q = Q^*$, so that the public debt ratio is determined by the no premium condition, $\frac{d}{TT^{\nu-1}y^H} \equiv \overline{d}$ or $d \equiv \overline{d}TT^{\nu-1}y^H$. Note also that since $Q = Q^*$, we set $\beta = \frac{1}{Q^*}$ in the parameterization stage.

In what follows, we present steady-state solutions of this economy. We will start with the case in which the country's net foreign debt position is unrestricted so that $\widetilde{f} \equiv \frac{(1-\lambda)TT^{1-\nu}d - TT^{\nu^*}f^h}{y^H}$ is endogenously determined. In turn, we will study the case in which we set the country's net foreign debt position equal to zero (meaning that the trade is balanced) so that $\widetilde{f} = 0$ in the new reformed steady state.

Steady State of the Reformed Economy with Unrestricted Foreign Debt

Now $Q = Q^*$ so that $d \equiv \overline{d}TT^{\nu-1}y^H$. The steady-state system is

$$1 = \beta[1 - \delta + (1 - \tau^k)r^k] \tag{110}$$

$$\chi_n n^\eta = c^{-\sigma}\frac{\nu c}{c^H}TT^{1-\nu}\frac{(1 - \tau_t^n)}{(1 + \tau_t^c)}w \tag{111}$$

$$\frac{c^H}{c^F} = \frac{\nu}{1 - \nu}TT \tag{112}$$

$$x = \delta k \tag{113}$$

$$c = \frac{\left(c^H\right)^\nu \left(c^F\right)^{1-\nu}}{(\nu)^\nu (1 - \nu)^{1-\nu}} \tag{114}$$

$$y^H = Ak^a n^{1-a} \tag{115}$$

$$y^H = c^H + x + s^g y^H + c^{F*} \tag{116}$$

$$r^k = TT^{1-v}mcaAk^{a-1}n^{1-a} \tag{117}$$

$$w = mc(1 - a)Ak^a n^{-a} \tag{118}$$

$$\widetilde{\omega} = \frac{1}{TT^{1-v}}y^H - \frac{1}{TT^{1-v}}r^k k - wn \tag{119}$$

$$z^1 = \frac{\phi}{(\phi - 1)}z^2 \tag{120}$$

Table A1. Steady-State Solution of the Reformed Economy with Unrestricted $\widetilde{f}$

Variables	Description	Steady-State Solution
u	Period Utility	0.9458
y^H	Output	0.858759
TT	Terms of Trade	0.838459
$Q - Q^*$	Interest Rate Premium	0
$TT^{1-\nu}\frac{c}{y^H}$	Consumption as Share of GDP	0.5522
$\frac{k}{y^H}$	Physical Capital as Share of GDP	4.2291
$\frac{TT^{\nu^*}f^h}{y^H}$	Private Foreign Assets as Share of GDP	−4.5742
$\frac{TT^{1-\nu}d}{y^H}$	Total Public Debt as Share of GDP	0.9
$\widetilde{f} \equiv \frac{(1-\lambda)dTT^{1-\nu} - TT^{\nu^*}f^h}{y^H}$	Total Foreign Debt as Share of GDP	4.8982

$$z^1 = yTT^{\nu-1} + \beta\theta z^1 \tag{121}$$

$$z^2 = ymc + \beta\theta z^2 \tag{122}$$

$$d \equiv \overline{d}TT^{\nu-1}y^H \tag{123}$$

$$\begin{aligned} d = Q^*d + TT^{\nu-1}s^g y^H - \tau^c\left(\frac{1}{TT^{1-v}}c^H + TT^v c^F\right) \\ - \tau^k\left(\frac{1}{TT^{1-v}}r^k k + \widetilde{\omega}\right) - \tau^n wn - TT^{\nu-1}s^l y_t^H \end{aligned} \tag{124}$$

$$\begin{aligned} (1-\lambda)d - TT^{\nu^*+\nu-1}f^h = -TT^{\nu-1}c^{F*} + TT^\nu c^F \\ + QTT^{\nu^*+\nu-1}\left(\frac{1}{TT^{v+v^*-1}}(1-\lambda)d - f^h\right). \end{aligned} \tag{125}$$

We thus have sixteen equations in c, c^H, c^F, k, x, n, y_t^H, TT, $mc, \widetilde{\omega}, d, z^1, z^2$, r^k, w, f^h. The numerical solution of this system is presented in table A1. Notice that, for the reasons explained in the text, private foreign debt, $-f^h > 0$, and hence the country's net foreign debt-to-GDP ratio, $\widetilde{f}$, are extremely high in this case.

Table A2. Steady-State Solution of the Reformed Economy with $\widetilde{f} = 0$ when the Residual Fiscal Instrument Is s^g

Variables	Description	Steady-State Solution
u	Period Utility	0.9125
r^k	Real Return to Physical Capital	0.0749
w	Real Wage Rate	1.2442
n	Hours Worked	0.3403
y^H	Output	0.8237
TT	Terms of Trade	1.01
s^g	Capital Tax Rate	0.2306
$Q - Q^*$	Interest Rate Premium	0
$TT^{1-\nu}\frac{c}{y^H}$	Consumption as Share of GDP	0.6002
$\frac{k}{y^H}$	Physical Capital as Share of GDP	4.2291
$TT^{\nu^*}\frac{f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.3240
$\frac{TT^{1-\nu}d}{y^H}$	Total Public Debt as Share of GDP	0.9
$\widetilde{f} \equiv \frac{(1-\lambda)dTT^{1-\nu}-TT^{\nu^*}f^h}{y^H}$	Total Foreign Debt as Share of GDP	0

Steady State of the Reformed Economy with Restricted Foreign Debt

Now $Q = Q^*$, so that $d \equiv \overline{d}TT^{\nu-1}y^H$ as above, and, in addition, $\widetilde{f} \equiv \frac{(1-\lambda)TT^{1-\nu}d-TT^{\nu^*}f^h}{y^H} = 0$, so that $(1-\lambda)\,dTT^{1-\nu} - TT^{\nu^*}f^h = 0$ or $f^h \equiv \frac{(1-\lambda)dTT^{1-\nu}}{TT^{\nu^*}}$. In this case, the equation for the balance of payments (125) simplifies to

$$TTc^F = c^{F*}. \tag{126}$$

Therefore, using this equation in the place of (125), we have sixteen equations in c, c^H, c^F, k, x, n, y_t^H, TT, mc, $\widetilde{\omega}, d, z^1, z^2$, r^k, w, and one of the tax-spending policy instruments, s^g, τ^c, τ^k, τ^n. Note that in turn f^h follows residually from $f^h \equiv \frac{(1-\lambda)dTT^{1-\nu}}{TT^{\nu^*}}$. The numerical solution of this system under each public financing instrument

Table A3. Steady-State Solution of the Reformed Economy with $\widetilde{f} = 0$ when the Residual Fiscal Instrument Is τ^c

Variables	Description	Steady-State Solution
u	Period Utility	0.9227
r^k	Real Return to Physical Capital	0.0749
w	Real Wage Rate	1.2442
n	Hours Worked	0.3403
y^H	Output	0.8237
TT	Terms of Trade	1.01
τ^c	Capital Tax Rate	0.1593
$Q - Q^*$	Interest Rate Premium	0
$TT^{1-\nu} \frac{c}{y^H}$	Consumption as Share of GDP	0.6086
$\frac{k}{y^H}$	Physical Capital as Share of GDP	4.2291
$TT^{\nu^*} \frac{f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.3240
$\frac{TT^{1-\nu} d}{y^H}$	Total Public Debt as Share of GDP	0.9
$\widetilde{f} \equiv \frac{(1-\lambda) d TT^{1-\nu} - TT^{\nu^*} f^h}{y^H}$	Total Foreign Debt as Share of GDP	0

is presented in tables A2–A5, while a summary of these solutions appears in table 3 in the main text.

Restricted Changes in Fiscal Policy Instruments

We now repeat the main computations restricting the magnitude of feedback coefficients in the policy rules so that all tax-spending policy instruments cannot change by more than 10 percentage points from their averages in the data (we report that the results also do not change when we allow for 5 percentage points only).

Restricted results for optimal feedback policy coefficients, the implied statistics, and welfare over various time horizons are reported in tables A6, A7, and A8, which correspond to tables 4, 5, and 6, respectively, in the main text.

Inspection of the new tables, and comparison to their counterparts in the main text, implies that the main qualitative results do

Table A4. Steady-State Solution of the Reformed Economy with $\widetilde{f} = 0$ when the Residual Fiscal Instrument Is τ^k

Variables	Description	Steady-State Solution
u	Period Utility	0.9311
r^k	Real Return to Physical Capital	0.0729
w	Real Wage Rate	1.2649
n	Hours Worked	0.3393
y^H	Output	0.8348
TT	Terms of Trade	1.01
τ^k	Capital Tax Rate	0.2930
$Q - Q^*$	Interest Rate Premium	0
$TT^{1-\nu} \frac{c}{y^H}$	Consumption as Share of GDP	0.6040
$\frac{k}{y^H}$	Physical Capital as Share of GDP	4.3448
$TT^{\nu^*} \frac{f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.3240
$\frac{TT^{1-\nu} d}{y^H}$	Total Public Debt as Share of GDP	0.9
$\widetilde{f} \equiv \frac{(1-\lambda) d TT^{1-\nu} - TT^{\nu^*} f^h}{y^H}$	Total Foreign Debt as Share of GDP	0

not change. Namely, debt consolidation is again preferable to non-debt consolidation after the first ten years. Also, although obviously feedback policy coefficients are now smaller, the best fiscal policy mix again implies that we should earmark public spending for the reduction of public debt and, at the same time, cut taxes to mitigate the recessionary effects of debt consolidation. The only difference is that now, since cuts in income taxes are restricted, we should also cut consumption taxes.

Impulse Response Functions (IRFs) with One Fiscal Instrument at a Time

Figures A1 and A2 compare debt consolidation results when the fiscal authorities can use all instruments jointly to the case in which they are restricted to use one instrument at a time only. Results are always with optimized rules. To save on space, we focus on results for the public debt-to-GDP ratio and private consumption.

Table A5. Steady-State Solution of the Reformed Economy with $\widetilde{f} = 0$ when the Residual Fiscal Instrument Is τ^n

Variables	Description	Steady-State Solution
u	Period Utility	0.9290
r^k	Real Return to Physical Capital	0.0749
w	Real Wage Rate	1.2442
n	Hours Worked	0.3435
y^H	Output	0.8314
TT	Terms of Trade	1.01
τ^n	Capital Tax Rate	0.4018
$Q - Q^*$	Interest Rate Premium	0
$TT^{1-\nu}\frac{c}{y^H}$	Consumption as Share of GDP	0.6086
$\frac{k}{y^H}$	Physical Capital as Share of GDP	4.2291
$TT^{\nu^*}\frac{f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.3240
$\frac{TT^{1-\nu}d}{y^H}$	Total Public Debt as Share of GDP	0.9
$\widetilde{f} \equiv \frac{(1-\lambda)dTT^{1-\nu} - TT^{\nu^*}f^h}{y^H}$	Total Foreign Debt as Share of GDP	0

Table A6. Optimal Reaction to Debt and Output with Debt Consolidation (restricted optimal fiscal policy mix)

Fiscal Instruments	Optimal Reaction to Debt	Optimal Reaction to Output
s_t^g	$\gamma_l^g = 0.4$	$\gamma_y^g = 0.0011$
τ_t^c	$\gamma_l^c = 0.0005$	$\gamma_y^c = 0.4865$
τ_t^k	$\gamma_l^k = 0.0026$	$\gamma_y^k = 0.9496$
τ_t^n	$\gamma_l^n = 0.0023$	$\gamma_y^n = 0.95$

Note: $V_0 = 79.9336$.

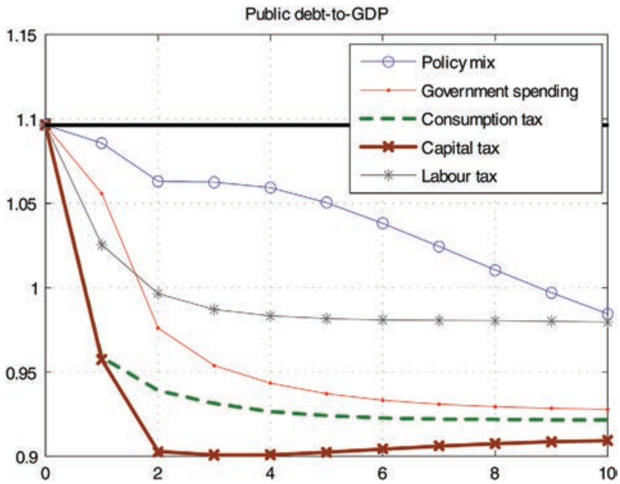
Table A7. Statistics Implied by Table A6

	Elasticity to Liabilities	Elasticity to Output	Min/Max	Five Periods Average	Ten Periods Average	Twenty Periods Average	Data Average
s_t^g	−2.016%	−0.0035%	0.1310/0.2273	0.1473	0.1576	0.1760	0.2222
τ_t^c	0.0032%	1.97%	0.1352/0.1799	0.1492	0.1545	0.1614	0.1756
τ_t^k	0.0093%	2.16%	0.2143/0.3014	0.2418	0.2520	0.2654	0.3118
τ_t^n	0.0062%	1.61%	0.3423/0.4295	0.3698	0.38	0.3934	0.421

Table A8. Welfare over Different Time Horizons with, and without, Debt Consolidation (restricted optimal fiscal policy mix)

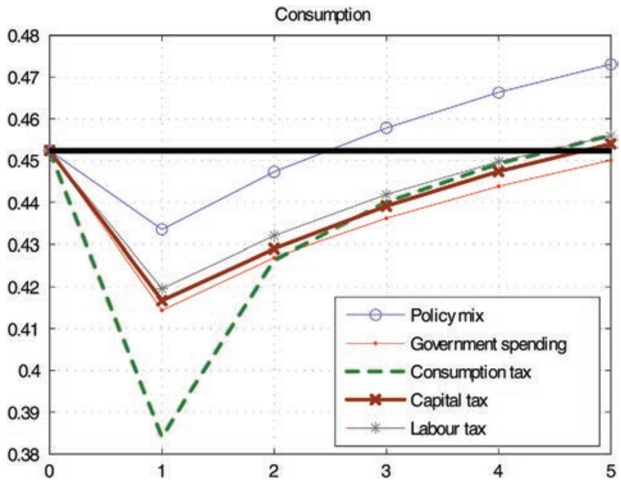
	Two Periods	Four Periods	Ten Periods	Twenty Periods	Thirty Periods
With Consolid.	1.4862	3.0164	7.7221	15.3294	22.2939
Without Consolid.	(1.6251)	(3.1791)	(7.4481)	(13.4127)	(18.1904)
Welfare Gain/Loss	−0.0458	−0.0327	0.0267	0.1075	0.1693

Figure A1. IRFs of Public Debt to GDP (comparison of alternative fiscal policies)



Note: IRFs are in levels and converge to the reformed steady state, while the solid horizontal line indicates the point of departure (status quo value).

Figure A2. IRFs of Consumption (comparison of alternative fiscal policies)



Note: IRFs are in levels and converge to the reformed steady state, while the solid horizontal line indicates the point of departure (status quo value).

Appendix 6. Robustness to the Public Debt Threshold Parameter

The Case with a Lower Public Debt Threshold

We now assume $\overline{d} \equiv 0.8$. This implies $\psi \equiv 0.0319$ to hit the data. Then, we have the solution shown in tables A9–A11.

Table A9. Status Quo Steady-State Solution

Variables	**Description**	**Steady-State Solution**	**Data**
u	Period Utility	0.8217	—
r^k	Real Return to Physical Capital	0.0908	—
w	Real Wage Rate	1.1140	—
n	Hours Worked	0.3313	0.2183
y^H	Output	0.7124	—
TT	Terms of Trade	0.9945	—
$Q - Q^*$	Interest Rate Premium	0	0.011
$TT^{1-\nu}\frac{c}{y^H}$	Consumption as Share of GDP	0.6334	0.5961
$\frac{k}{y^H}$	Physical Capital as Share of GDP	3.4872	—
$TT^{\nu^*}\frac{f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.1749	0.1039
$\frac{TT^{1-\nu}d}{y^H}$	Total Public Debt as Share of GDP	1.0965	1.0828
$\widetilde{f} \equiv \frac{(1-\lambda)dTT^{1-\nu} - TT^{\nu^*}f^h}{y^H}$	Total Foreign Debt as Share of GDP	0.2194	0.2109

Table A10. Reformed Steady-State Solution

Variables	Description	Steady-State Solution
u	Period Utility	0.9326
r^k	Real Return to Physical Capital	0.0727
w	Real Wage Rate	1.2673
n	Hours Worked	0.3394
y^H	Output	0.8367
TT	Terms of Trade	1.01
τ^k	Capital Tax Rate	0.2908
$Q - Q^*$	Interest Rate Premium	0
$TT^{1-\nu} \frac{c}{y^H}$	Consumption as Share of GDP	0.6035
$\frac{k}{y^H}$	Physical Capital as Share of GDP	4.3583
$TT^{\nu^*} \frac{f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.2880
$\frac{TT^{1-\nu} d}{y^H}$	Total Public Debt as Share of GDP	0.8
$\widetilde{f} \equiv \frac{(1-\lambda) d TT^{1-\nu} - TT^{\nu^*} f^h}{y^H}$	Total Foreign Debt as Share of GDP	0

Table A11. Optimal Reaction to Debt and Output with Debt Consolidation (optimal fiscal policy mix)

Fiscal Instruments	Optimal Reaction to Debt	Optimal Reaction to Output
s_t^g	$\gamma_l^g = 0.4338$	$\gamma_y^g = 0$
τ_t^c	$\gamma_l^c = 0$	$\gamma_y^c = 0.0279$
τ_t^k	$\gamma_l^k = 0$	$\gamma_y^k = 2.2569$
τ_t^n	$\gamma_l^n = 0.0001$	$\gamma_y^n = 2.136$
Note: $V_0 = 80.1038$.		

The Case with a Higher Public Debt Threshold

Next, we assume $\overline{d} \equiv 1$. This implies $\psi \equiv 0.108$ to hit the data. Then, we have the solution shown in tables A12–A14.

Table A12. Status Quo Steady-State Solution

Variables	Description	Steady-State Solution	Data
u	Period Utility	0.8217	—
r^k	Real Return to Physical Capital	0.0908092	—
w	Real Wage Rate	1.11373	—
n	Hours Worked	0.331273	0.2183
y^H	Output	0.712311	—
TT	Terms of Trade	0.995009	—
$Q - Q^*$	Interest Rate Premium	0.011	0.011
$TT^{1-\nu}\frac{c}{y^H}$	Consumption as Share of GDP	0.6335	0.5961
$\frac{k}{y^H}$	Physical Capital as Share of GDP	3.4872	—
$TT^{\nu^*}\frac{f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.1827	0.1039
$\frac{TT^{1-\nu}d}{y^H}$	Total Public Debt as Share of GDP	1.0965	1.0828
$\widetilde{f} \equiv \frac{(1-\lambda)dTT^{1-\nu} - TT^{\nu^*}f^h}{y^H}$	Total Foreign Debt as Share of GDP	0.2122	0.2109

Table A13. Reformed Steady-State Solution

Variables	Description	Steady-State Solution
u	Period Utility	0.9296
r^k	Real Return to Physical Capital	0.0731
w	Real Wage Rate	1.2625
n	Hours Worked	0.3391
y^H	Output	0.8328
TT	Terms of Trade	1.01
τ^k	Capital Tax Rate	0.2951
$Q - Q^*$	Interest Rate Premium	0
$TT^{1-\nu} \frac{c}{y^H}$	Consumption as Share of GDP	0.6045
$\frac{k}{y^H}$	Physical Capital as Share of GDP	4.3314
$TT^{\nu^*} \frac{f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.3600
$\frac{TT^{1-\nu} d}{y^H}$	Total Public Debt as Share of GDP	1
$\widetilde{f} \equiv \frac{(1-\lambda) d TT^{1-\nu} - TT^{\nu^*} f^h}{y^H}$	Total Foreign Debt as Share of GDP	0

Table A14. Optimal Reaction to Debt and Output with Debt Consolidation (optimal fiscal policy mix)

Fiscal Instruments	Optimal Reaction to Debt	Optimal Reaction to Output
s_t^g	$\gamma_l^g = 0.4588$	$\gamma_y^g = 0.0017$
τ_t^c	$\gamma_l^c = 0$	$\gamma_y^c = 0.0018$
τ_t^k	$\gamma_l^k = 0.0014$	$\gamma_y^k = 2.2569$
τ_t^n	$\gamma_l^n = 0.1306$	$\gamma_y^n = 1.5629$

Note: $V_0 = 79.8482$.

Appendix 7. Robustness to the Net Foreign Debt Position in the Reformed Steady State

The Case with 0.1 Net Foreign Debt Ratio in the Reformed Steady State

When we set $\widetilde{f} \equiv \frac{(1-\lambda)TT^{1-\nu}d - TT^{\nu^*}f^h}{y^H} = 0.1$ in the reformed steady state, the results are as shown in tables A15 and A16.

Table A15. Steady-State Solution of the Reformed Economy with $\widetilde{f} = 0.1$

Variables	Description	Steady-State Solution
u	Period Utility	0.9315
y^H	Output	0.835228
TT	Terms of Trade	1.00615
$Q - Q^*$	Interest Rate Premium	0
$TT^{1-\nu}\frac{c}{y^H}$	Consumption as Share of GDP	0.6029
$\frac{k}{y^H}$	Physical Capital as Share of GDP	4.3425
$TT^{\nu^*}\frac{f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.2240
$\frac{TT^{1-\nu}d}{y^H}$	Total Public Debt as Share of GDP	0.9
$\widetilde{f} \equiv \frac{(1-\lambda)dTT^{1-\nu} - TT^{\nu^*}f^h}{y^H}$	Total Foreign Debt as Share of GDP	0.1

Table A16. Optimal Reaction to Debt and Output with Debt Consolidation (optimal fiscal policy mix)

Fiscal Instruments	Optimal Reaction to Debt	Optimal Reaction to Output
s_t^g	$\gamma_l^g = 0.6725$	$\gamma_y^g = 0.0020$
τ_t^c	$\gamma_l^c = 0$	$\gamma_y^c = 0.7435$
τ_t^k	$\gamma_l^k = 0.0011$	$\gamma_y^k = 2.2572$
τ_t^n	$\gamma_l^n = 0.0015$	$\gamma_y^n = 2.14$

Note: $V_0 = 80.2538$.

The Case with 0.2109 Net Foreign Debt Ratio in the Reformed Steady State

When we set $\widetilde{f} \equiv \frac{(1-\lambda)TT^{1-\nu}d - TT^{\nu^*}f^h}{y^H} = 0.2109$ (as in the data) in the reformed steady state, the results are as shown in tables A17 and A18.

Table A17. Steady-State Solution of the Reformed Economy with $\widetilde{f} = 0.2109$

Variables	Description	Steady-State Solution
u	Period Utility	0.9320
y^H	Output	0.8358
TT	Terms of Trade	1.002
$Q - Q^*$	Interest Rate Premium	0
$TT^{1-\nu}\frac{c}{y^H}$	Consumption as Share of GDP	0.6017
$\frac{k}{y^H}$	Physical Capital as Share of GDP	4.3397
$TT^{\nu^*}\frac{f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.1050
$\frac{TT^{1-\nu}d}{y^H}$	Total Public Debt as Share of GDP	0.9
$\widetilde{f} \equiv \frac{(1-\lambda)dTT^{1-\nu} - TT^{\nu^*}f^h}{y^H}$	Total Foreign Debt as Share of GDP	0.2109

Table A18. Optimal Reaction to Debt and Output with Debt Consolidation (optimal fiscal policy mix)

Fiscal Instruments	Optimal Reaction to Debt	Optimal Reaction to Output
s_t^g	$\gamma_l^g = 0.5558$	$\gamma_y^g = 0.0114$
τ_t^c	$\gamma_l^c = 0.0094$	$\gamma_y^c = 1.3091$
τ_t^k	$\gamma_l^k = 0.0159$	$\gamma_y^k = 1.3759$
τ_t^n	$\gamma_l^n = 0$	$\gamma_y^n = 2.14$

Note: $V_0 = 80.3931$.

Appendix 8. Robustness to World Interest Rate Shocks

Using the new specification presented in the main text, the results are as shown in tables A19 and A20.

Table A19. Optimal Reaction to Debt and Output with Debt Consolidation (optimal fiscal policy mix)

Fiscal Instruments	Optimal Reaction to Debt	Optimal Reaction to Output
s_t^g	$\gamma_l^g = 0.2068$	$\gamma_y^g = 0$
τ_t^c	$\gamma_l^c = 0.0002$	$\gamma_y^c = 0.2097$
τ_t^k	$\gamma_l^k = 0$	$\gamma_y^k = 2.2569$
τ_t^n	$\gamma_l^n = 0$	$\gamma_y^n = 2.1360$

Note: $V_0 = 79.5457$.

Table A20. Welfare over Different Time Horizons with, and without, Debt Consolidation (optimal fiscal policy mix)

	Two Periods	Four Periods	Ten Periods	Twenty Periods	Thirty Periods
With Consolid.	1.4589	2.8359	7.0660	14.4129	21.9026
Without Consolid.[a]	(1.5236)	(2.9533)	(6.8521)	(12.6868)	(17.8657)
Welfare Gain/Loss	−0.0216	−0.0237	0.0208	0.0963	0.1662

[a]The values of the optimized feedback policy coefficients when we start from, and return to, the same status quo steady state are $\gamma_l^g = 0.0109, \gamma_l^c = 0.5206, \gamma_l^k = 0, \gamma_l^n = 0.3032, \gamma_y^g = 0.0129, \gamma_y^c = 0, \gamma_y^k = 2.2564$, and $\gamma_y^n = 0.0236$.

Appendix 9. Ramsey Policy Problem

The Ramsey government chooses the paths of policy instruments to maximize the household's expected discounted lifetime utility, V_t, subject to the equations of the DE. As shown in appendix 4 above, the DE system contains twenty-four equations (we do not include the equation for V_t).

As said in the main text, we solve for a "timeless" Ramsey equilibrium, meaning that the resulting equilibrium conditions are time invariant (see also, e.g., Schmitt-Grohé and Uribe 2005). If we follow the dual approach to the Ramsey policy problem, the government chooses the paths of $\{y_t^H, c_t, c_t^H, c_t^F, n_t, x_t, k_t, f_t^h, TT_t, \Pi_t, \Pi_t^H, \Theta_t, \Delta_t, w_t, mc_t, \widetilde{\omega}_t, r_t^k, Q_t, d_t, \Pi_t^*, z_t^1, z_t^2, R_t, l_t\}_{t=0}^{\infty}$, which are the twenty-four endogenous variables of the DE system, plus the paths of the three tax rates, $\{\tau_t^c, \tau_t^k, \tau_t^n\}_{t=0}^{\infty}$. Notice that public debt, $\{d_t\}_{t=0}^{\infty}$, is also among the choice variables. Also notice that, in solving this optimization problem, we take as given the rate of exchange rate depreciation, $\{\epsilon_t\}_{t=0}^{\infty}$; public spending on goods/services, $\{s_t^g\}_{t=0}^{\infty}$; lump-sum government transfers, $\{s_t^l\}_{t=0}^{\infty}$; and the share of public debt issued for the domestic market, $\{\lambda_t\}_{t=0}^{\infty}$. We treat these variables as exogenous for different reasons: in a currency union model, $\epsilon_t = 1$ all the time; we take $\{s_t^g\}_{t=0}^{\infty}$ as given for simplicity; we take $\{s_t^l\}_{t=0}^{\infty}$ as given because, if the Ramsey government can choose a lump-sum policy instrument, then its problem becomes trivial; finally, we set λ_t exogenously and equal to its data average value (as we have done so far anyway) because, if it is chosen by the Ramsey government, its optimality condition will be identical to the household's optimality condition for financial assets/debt in a Ramsey steady state without sovereign interest rate premia, so the Ramsey steady-state system will be indeterminate (in particular, in the steady state, the government's optimality condition for λ_t is reduced to $1 = \beta Q$, which is also the household's optimality condition for foreign assets/debt). For the very same reason (namely, the optimality conditions of the household and the government with respect to f_t^h are both reduced to $1 = \beta Q$ in the Ramsey steady state without sovereign interest rate premia), we set f_t^h so as to satisfy $\widetilde{f} = 0$ in the reformed steady state (as we have done so far anyway in the reformed steady state). Thus, from $\widetilde{f} \equiv \frac{(1-\lambda)dTT^{1-\nu} - TT^{\nu^*} f^h}{y^H} = 0$, it follows $f^h = (1-\lambda)\,\overline{d}TT^{\nu-1}y^H$.

Table A21. Ramsey Steady-State Solution (unrestricted)

Variables	Description	Steady-State Solution
u	Period Utility	1.2215
y^H	Output	1.56176
TT	Terms of Trade	1.01
τ^c	"Optimal" Consumption Tax Rate	26.5726
τ^k	"Optimal" Capital Tax Rate	0.00272056
τ^n	"Optimal" Consumption Tax Rate	−26.4976
$Q - Q^*$	Interest Rate Premium	0
$TT^{1-\nu} \frac{c}{y^H}$	Consumption as Share of GDP	0.5327
$\frac{k}{y^H}$	Physical Capital as Share of GDP	6.1284
$TT^{\nu^*} \frac{f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.3240
$\frac{TT^{1-\nu} d}{y^H}$	Total Public Debt as Share of GDP	0.9
$\widetilde{f} \equiv \frac{(1-\lambda) d TT^{1-\nu} - TT^{\nu^*} f^h}{y^H}$	Total Foreign Debt as Share of GDP	0

Working in this way, and including Ramsey multipliers, we end up with a well-defined system consisting of fifty-one equations in fifty-one endogenous variables (the system is available upon request from the authors). The numerical solution of this system in the steady state is presented in table A21 (we do not report solutions for Ramsey multipliers). In this solution, as said before, $Q = Q^* = R = \frac{1}{\beta}$, and $f^h = (1-\lambda)\,\overline{d} TT^{\nu-1} y^H$. (Note that nominal variables, like R, do not affect the real allocation, and hence utility, at steady state.)

Notice, as also said in the text, and as is well established in the Ramsey literature with optimally chosen consumption taxes, that the solution in table A21 is characterized by an extremely large consumption tax rate, $\tau^c = 26.5726$, and an equally large labor

Table A22. Ramsey Steady-State Solution (restricted $\tau_t^n = 0$)

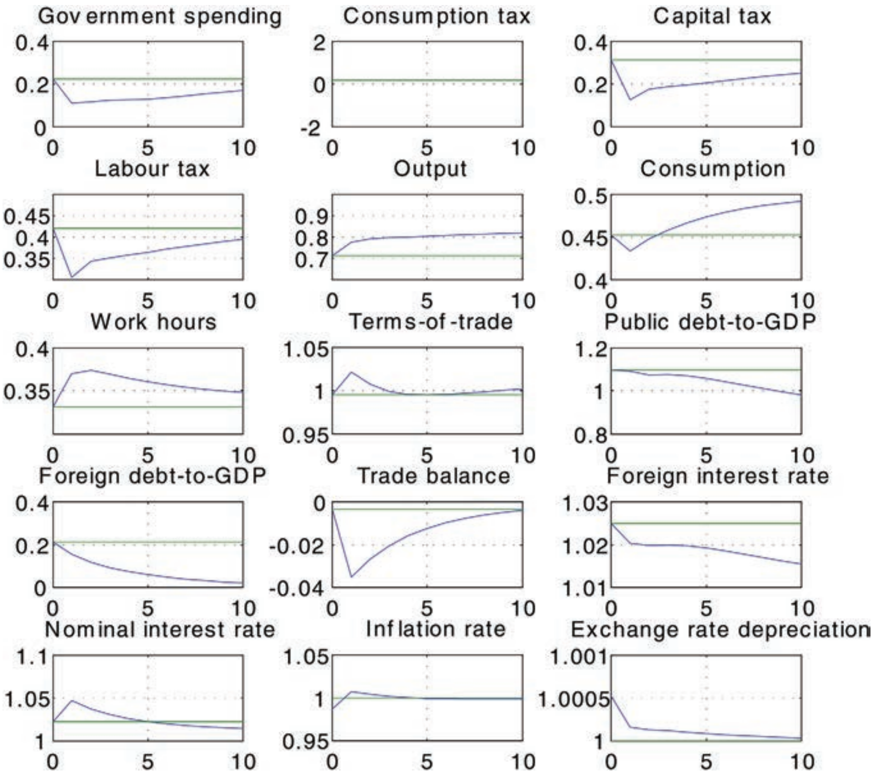
Variables	Description	Steady-State Solution
u	Period Utility	1.0950
y^H	Output	1.11622
TT	Terms of Trade	1.01
τ^k	"Optimal" Capital Tax Rate	0.0500465
τ^c	"Optimal" Consumption Tax Rate	0.810199
$Q - Q^*$	Interest Rate Premium	0
$TT^{1-\nu} \frac{c}{y^H}$	Consumption as Share of GDP	0.5443
$\frac{k}{y^H}$	Physical Capital as Share of GDP	5.8387
$TT^{\nu^*} \frac{f^h}{y^H}$	Private Foreign Assets as Share of GDP	0.3240
$\frac{TT^{1-\nu} d}{y^H}$	Total Public Debt as Share of GDP	0.9
$\widetilde{f} \equiv \frac{(1-\lambda) d TT^{1-\nu} - TT^{\nu^*} f^h}{y^H}$	Total Foreign Debt as Share of GDP	0

subsidy, $\tau^n = -26.4976 < 0$. Hence, following the same literature, we rule out a subsidy to labor by setting $\tau_t^n = 0$ at all t. This gives the solution in table A22 (this table is also presented in the main text, but we repeat it here for the reader's convenience).

Appendix 10. Flexible Exchange Rates

The IRFs under flexible exchange rates are shown in figure A3 (the zero lower bound for the nominal interest rate is not violated).

Figure A3. IRFs under Flexible Exchange Rates



Note: IRFs are in levels and converge to the reformed steady state, while the solid horizontal line indicates the point of departure (status quo value).

References

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): 926–48.

Rabanal, P. 2009. "Inflation Differentials between Spain and the EMU: A DSGE Perspective." *Journal of Money, Credit and Banking* 41 (6): 1141–66.

Schmitt-Grohé, S., and M. Uribe. 2004. "Optimal Fiscal and Monetary Policy under Sticky Prices." *Journal of Economic Theory* 114 (2): 198–230.

———. 2005. "Optimal Fiscal and Monetary Policy in a Medium-Scale Macroeconomic Model." In *NBER Macroeconomics Annual,* ed. M. Gertler and K. Rogoff, 385–425. Cambridge, MA: MIT Press.

———. 2007. "Optimal Simple and Implementable Monetary and Fiscal Rules." *Journal of Monetary Economics* 54 (6): 1702–25.