Volume 15, Number 4 October 2019

IJCB INTERNATIONAL JOURNAL OF CENTRAL BANKING Volume 15, Number 4 October 2019

IJCB
INTERNATIONAL JOURNAL OF CENTRAL BANKING

International Journal of Central Banking

Volume 15, Number 4 October 2019

*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 (001 216 579 3125), via fax (001 216 579 2886), 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
Executive Office Department
Federal Reserve Bank of Cleveland
P.O. Box 6387
Cleveland, OH 44101-1387
USA

Phone: 001 216 579 3125
Fax: 001 216 579 2886
E-mail: editor@ijcb.org

ISSN: 1815-4654

Budget-Neutral Labor Tax Wedge Reductions: A Simulation-Based Analysis for the Euro Area*

Maria-Grazia Attinasi,[a] Doris Prammer,[b] Nikolai Stähler,[c] Martino Tasso,[d] and Stefan van Parys[e]
[a]European Central Bank
[b]Oesterreichische Nationalbank
[c]Deutsche Bundesbank
[d]Banca d'Italia
[e]National Bank of Belgium

Budget-neutral tax wedge reductions rank high in the policy agenda of several EMU member states. Using a New Keynesian DSGE model of a monetary union with a search-and-matching market structure and a fiscal bloc containing a wide range of taxes and disaggregated government spending, we evaluate the macroeconomic and welfare effects of reducing the firms' and workers' labor tax rates under alternative financing instruments. Overall, a tax wedge reduction is beneficial in terms of both welfare and output. While financing the labor tax wedge reduction by an increase in consumption taxation yields most favorable output effects, financing it by a reduction in government spending is more welfare enhancing, as the latter does not imply a policy-induced increase in private consumption costs. We also show that, when there exists an extensive and intensive labor margin, a reduction in the workers' and not the firms' burden can be most beneficial.

JEL Codes: H2, J6, E32, E62.

*The authors thank Othman Bouabdallah, Giovanni Callegari, Francesco Drudi, Doug Hamilton, Mathias Hoffmann, Christophe Kamps, Ralph Lütticke, Jochen Mankart, Javier Pérez, Thomas Warmedinger, members of the European System of Central Banks' Working Group on Public Finances (WGPF), participants of the 2015 National Tax Association Annual Conference in Boston, and participants of the 2017 Annual Meeting of the Austrian Economics Society for helpful comments. Nikolai Stähler gratefully acknowledges the hospitality of the ECB's Fiscal Policy Division, where a significant part of this work has been undertaken. The paper represents the authors' personal opinions and does not necessarily reflect the views of their institutions. Corresponding author e-mail: nikolai.staehler@bundesbank.de.

1. Introduction

In most euro-area countries, the tax wedge on labor income is large relative to international standards (European Commission 2014). Since 2011, the European Commission's country-specific recommendations include calls for reducing the tax wedge. In July 2014, the Eurogroup identified lowering labor taxes as a top policy priority. Moreover, given the lack of fiscal space in many member countries, it was recommended that such reforms should be implemented in a budget-neutral way. The issue is again addressed in a recent publication (European Commission 2016). Furthermore, given the weak labor market performance in the aftermath of the crisis and the fact that, within the euro zone, about 18 million people are unemployed (yielding an average unemployment rate of 11 percent),[1] it is even claimed that "the crisis will only be over when unemployment falls to socially sustainable levels" and that "jobs fail to be created ... not because of the 'lack of demand' as often claimed, but mainly because wage costs are high relative to productivity [and] social insurance and tax burdens are heavy" (Thimann 2015).

This paper analyzes the macroeconomic and welfare effects of a budget-neutral reduction in the labor tax wedge. We use a macroeconomic New Keynesian dynamic stochastic general equilibrium (DSGE) model of a monetary union with a search-and-matching labor market and a fiscal bloc containing a wide range of taxes and government spending. The model-based framework allows analysis of the issue at hand from different angles. First, we explore how different ways of achieving budget neutrality (e.g., higher consumption taxes versus lower government spending or employment) affect the economy. Second, we assess whether cutting social security contributions or personal income taxes matters for these effects, especially in the presence of an intensive labor margin on top of the extensive one. Finally, we test for the role of country-specific characteristics (e.g., trade openness, overall labor market efficiency, and country size).

[1] See http://ec.europa.eu/eurostat/statistics-explained/index.php/Unemployment_statistics# Unemployment_trends.

To the best of our knowledge, this is the first paper to model the effects of a budget-neutral reduction in the labor tax wedge in the context of a search-and-matching labor market including an extensive (i.e., whether to hire a new worker or not) and intensive (i.e., deciding how many hours to work) labor adjustment margin. This labor market structure allows for a better understanding of how cuts in social security contributions or personal income taxes affect labor demand and supply, along with the role of structural labor market features (for example, the bargaining process).

A budget-neutral reduction in the labor tax wedge is found to have positive macroeconomic effects. The larger gains are associated with a fiscal devaluation—that is, a cut in the tax wedge financed by an increase in consumption taxes as defined in Engler et al. (2017)[2]—as opposed to a reduction in government purchases as a means to finance a lower tax wedge. This is because in our model government purchases contain a full home bias, while part of private demand is spent on imported goods. Hence, the increase in private demand stemming from a higher net labor income cannot fully compensate for the decline in aggregate demand (caused by the cut in government purchases). Findings are reversed in terms of households' welfare. Unlike a fiscal devaluation, decreasing public purchases does not depress private consumption via the policy-induced increase in consumption costs.

When an intensive labor margin is included in the model (i.e., firms and workers bargain over both wages and hours), financing a tax wedge cut via lower income taxes yields higher output gains than cutting firms' social security contributions. The reason is that lower income taxes immediately translate into higher net wages, hence

[2]In a strict sense, fiscal devaluation is defined as an intended nominal devaluation that can be robustly replicated with a small set of fiscal instruments: lower labor income taxes financed by a higher value-added tax (VAT) rate, where there is a sequence of taxes that replicates a sequence of nominal exchange rates while leaving the labor tax wedge constant (see Farhi, Gopinath, and Itskhoki 2014 and Kaufmann 2016). The literature, which we discuss in more detail below, however, also defines a permanent shift from labor income taxes to consumption taxes, where the latter is typically used as a proxy for VAT in New Keynesian models, as fiscal devaluation. We follow this definition in the paper. Also considering the similarities between a VAT and a generic consumption tax, qualitatively our results should be unaffected by this choice.

generating a higher disposable income. Firms, on their side, have an incentive to adjust labor input via the intensive margin rather than hiring new workers, as they save on search costs. Hiring new workers becomes more attractive when firms' tax burden is reduced (i.e., the cut in the tax wedge is implemented via lower social security contributions). In this case, the increase in workers' net wage income is relatively lower, as it is only indirectly affected by the measure. Workers are then less willing to supply more hours of work, which firms partly compensate for by increasing job creation. The additional search costs lead to a smaller increase in available net income so that the increase in private consumption is smaller than a reduction in the personal income tax. It should be noted that in the absence of an intensive labor margin (i.e., when hours worked are kept at the initial steady-state value), our model predicts that a cut in social security contributions yields higher output gains than a cut in the personal income tax, as the direct reduction in labor costs leads to a stronger fall in unemployment.

We find that country-specific characteristics matter for the model results in quantitative terms, but not qualitatively. Furthermore, spillovers to the rest of the euro area are positive, small, and depend on country size.

The rest of the paper is organized as follows. In section 2, we discuss related literature. Section 3 describes the model and its calibration. In section 4, we present the simulation design, while section 5 discusses the results. A welfare assessment can be found in section 6. Section 7 concludes.

2. Related Literature

The impact of taxes on the labor market has been addressed from several angles in the economic literature. Empirical macroeconomic studies mostly use aggregate data and perform cross-country comparisons and, in line with microeconomic theory (see Meghir and Phillips 2010, and Keane 2011), they usually find harmful effects of tax wedges on employment (e.g., Daveri and Tabellini 2000, and Bassanini and Duval 2006). By calibrating a simple labor supply model to the features of the United States and the main European economies, Prescott (2004) finds that the differences in aggregated hours of work across the Atlantic are primarily driven by observed discrepancies in marginal effective tax rates.[3]

Coenen, McAdam, and Straub (2008) analyze Prescott's insight through the lenses of a DSGE model. They find that reducing European tax wedges to levels comparable to the ones prevailing in the United States would increase the number of total hours worked by about 10 percent and significantly boost GDP in the long run. Ohanian, Raffo, and Rogerson (2008), using the framework of a neoclassical growth model calibrated to the economies of OECD countries over 1956–2004, also find that changes in tax rates explain most of the variability in worked hours across countries and through time.

Reductions in the labor tax wedge financed by higher consumption taxation have recently also been discussed with a focus on international competitiveness, often referred to as fiscal devaluation. Farhi, Gopinath, and Itskhoki (2014) provide a formal analysis of fiscal devaluations in a New Keynesian open economy DSGE model. They find that an intended nominal devaluation can be robustly replicated with a small set of fiscal instruments (namely labor income and consumption taxes). However, their contribution also shows that one should not over-estimate fiscal devaluation as a policy tool, as it may require substantial changes in tax rates. For example, a 10 percent nominal devaluation in Spain would require an increase of VAT taxes of as much as 7.6 percentage points.

Gadatsch, Stähler, and Weigert (2016) show that Germany's fiscal devaluation from 1999 to 2003 (generating a decrease in effective labor taxation by about 2 percentage points) improved GDP by only about one-quarter percentage point. Similarly, Lipinska and von Thadden (2009) show in a two-country DSGE model with a Walrasian labor market without matching frictions that fiscal devaluations generate only small quantitative effects. Stähler and Thomas (2012) and Boscá, Doménech, and Ferri (2013) show positive effects of fiscal devaluation in Spain. The positive effect of a fiscal devaluation is also confirmed by Gomes, Jacquinot, and Pisani (2016),

[3]This paper spurred a long series of reactions. Alesina, Glaeser, and Sacerdote (2006) present a critical evaluation of Prescott's argument. Even though the authors recognize the importance of taxes, they consider other labor market institutions more relevant. Empirically, Nickell, Nunziata, and Ochel (2005) discuss this issue for OECD countries.

who include Portugal in their analysis, and CPB (2013), the latter using country-specific general equilibrium models for four euro-area countries. Using the multi-country version of the Commission's QUEST model, the European Commission (2013) shows that fiscal devaluation mildly affects GDP positively already in the short to medium run, and can indeed significantly increase GDP in the long run. They also compare targeted tax reductions for differently skilled workers and find that cutting labor income tax rates for low-skilled (and, therefore, low-wage-earning) workers further augments GDP improvements because these workers exhibit a higher labor supply elasticity. Langot, Patureau, and Sopraseuth (2014) also find beneficial effects of fiscal devaluation in a model-based analysis for France.

The existing literature on fiscal devaluation usually focuses on one type of fiscal devaluation, i.e., either a reduction in employees' labor taxation or a reduction in employers' social security contribution. Notable exceptions are Burgert and Roeger (2014) and Engler et al. (2017). Engler et al. (2017) show that if only employers' social security contributions are decreased (instead of employees' and employers' contributions or labor taxes per se as done in the similar model by Lipinska and von Thadden 2009), the expected effects can be somewhat larger, which they attribute to higher competitiveness gains. Burgert and Roeger (2014) assess the efficiency of both types of fiscal devaluation using the European Commission's QUEST III model. They conclude that the long-run effects are identical in both scenarios; only the short-term efficiency is higher if employees' labor taxes are reduced. To our knowledge, there are no studies simulating fiscal devaluations with DSGE models that simultaneously incorporate a search-and-matching labor market and an intensive and an extensive labor adjustment margin, as we have in this study. The search-and-matching labor market together with two labor adjustment margins explains why there is a difference between cutting personal labor income tax or social security contribution rates primarily through the bargaining process. Our paper shows that, when ignoring the hours margin, a cut in social security contributions dominates a cut in the personal labor income tax rate, in line with the argumentation of Engler et al. (2017). The opposite holds when taking into account the hours margin, which is a result of the stronger increase in available net labor income and a positive wealth effect.

Distributional effects of fiscal devaluation have not gained much attention in the theoretical literature so far. In a micro-simulation study, CPB (2013) and Picos-Sánchez and Thomas (2015) find that fiscal devaluation tends to be regressive. This is confirmed by Burgert and Roeger (2014), especially in a situation in which transfer income recipients are not compensated for the increase in consumption taxes. The CPB (2013) qualifies, however, that if the cuts to social security contributions are targeted to low-income earners, a fiscal devaluation becomes progressive.

3. The Economic Environment

This section first describes the model used in the analysis and then turns to its calibration.

3.1 The Model

Overall, the model is quite a prototypical New Keynesian DSGE model in line with Smets and Wouters (2003, 2007), Christiano, Eichenbaum, and Evans (2005), Christoffel et al. (2009), and Boscá, Doménech, and Ferri (2011) and closely related to Stähler and Thomas (2012). It features a two-country monetary union structure. The integration of a labor market with search characteristics, based on Pissarides (2000), allows the inclusion of involuntary unemployment. We also introduce an intensive margin of labor supply in the model, meaning that labor supply and demand are fully flexible in number of hours. Furthermore, the model also contains a comprehensive public sector. The latter two elements make the model particularly well suited for the purpose of our analysis.

For what follows, we normalize population size of the entire European monetary union to unity, of which $\omega \in (0, 1)$ live in Home (labeled as the periphery), while the remaining $(1 - \omega)$ live in the rest of the EMU (labeled as the core). Throughout the formal model description, quantity variables will be expressed in per capita terms, unless otherwise indicated. Both regions are modeled analogously, while we allow structural parameters to differ. Hence, we restrict ourselves to explaining in detail only the home country. If the explicit description of the foreign country is necessary, we use asterisks to denote decisions made by the corresponding foreign agents as well as the structural parameters.

3.1.1 Households

As in Galí, Lopez-Salido, and Vallés (2007), each country is populated by a share $(1-\mu)$ of Ricardian households who have access to capital markets and, therefore, substitute consumption intertemporally (optimizers). The remaining share $\mu \in [0,1)$ is considered to be liquidity constrained in the sense that they consume all their labor income in each period ("rule-of-thumb," or RoT, household). Each type of household at time $t=0$ maximizes

$$\mathcal{W}_0^i = E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \cdot \underbrace{\left(\frac{\left(c_t^i - h \cdot c_{t-1}^i\right)^{1-\sigma_c}}{1-\sigma_c} - \kappa^h \cdot \left(n_t^{i,p} \cdot \frac{{l_t^{i,p}}^{1+\sigma_h}}{1+\sigma_h} + n_t^{i,g} \cdot \frac{{l_t^{i,g}}^{1+\sigma_h}}{1+\sigma_h} \right) \right)}_{=U(c^i, l^{i,p}, l^{i,g})} \right\}, \tag{1}$$

where E_t is the expectations operator conditional on time-t information, c_t^i denotes household consumption of final goods, and the superscripts $i = o, r$ denote optimizing and RoT households, respectively. h denotes the degree of habit formation in consumption.

Inside each household, its members may be employed in the public sector (denoted by $n_t^{i,g}$), employed in the private sector (denoted by $n_t^{i,p}$), or unemployed (denoted by u_t^i), with n expressed in terms of number of hours worked. Households face disutility of providing hours worked, $l_t^{i,p}$ or $l_t^{i,g}$, once employed in one of the two sectors. Disutility increases in the number of hours worked, where κ^h is a scaling parameter and σ_h a shape parameter of the disutility function. As becomes clear below, we will assume full consumption insurance within each household type as in Merz (1995) or Andolfatto (1996).

Households in both countries trade consumption and investment goods as well as international nominal bonds. The consumption and investment baskets, c_t^i and I_t^o, respectively, of a household of type i (only type o for investment) in the home country are given by

$$x_t^i = \left(\frac{x_{At}^i}{\vartheta} \right)^{\vartheta} \left(\frac{x_{Bt}^i}{1-\vartheta} \right)^{1-\vartheta},$$

with $x_t^i = \{c_t^i, I_t^o\}$, where c_{At}^i, I_{At}^o and c_{Bt}^i, I_{Bt}^o represent consumption/investment demand of goods produced in Home (country A) and Foreign (region B), respectively, and $\vartheta = \omega + \psi$ measures the preference for region-A goods, where ψ is a parameter capturing the degree of home bias in private consumption/investment. Note that $\vartheta^* = \omega - \psi^*$. From now onwards, let $p_{Bt} \equiv P_{Bt}/P_{At}$ denote the *terms of trade*, where P_{At} and P_{Bt} are the *producer price indexes* (PPIs) in countries A and B, respectively. Cost minimization by the household then implies $x_{At}^i/x_{Bt}^i = (\vartheta)/(1-\vartheta) \cdot p_{Bt}$. Nominal expenditure in consumption and investment goods equal $P_{At}c_{At}^i + P_{Bt}c_{Bt}^i = P_t c_t^i$ and $P_{At}I_{At}^o + P_{Bt}I_{Bt}^o = P_t I_t^o$, respectively, where $P_t = (P_{At})^{\vartheta}(P_{Bt})^{1-\vartheta}$ is the corresponding *consumer price index* (CPI). Notice that $P_t = P_{At} \cdot p_{Bt}^{1-\vartheta}$. Foreign-country CPI is analogously given by $P_t^* = P_{At}^{\vartheta^*} P_{Bt}^{1-\vartheta^*} = P_{Bt}(1/p_{Bt})^{\vartheta^*}$. Therefore, CPI inflation, $\pi_t \equiv P_t/P_{t-1}$, evolves according to $\pi_t = \pi_{At}(p_{Bt}/p_{Bt-1})^{1-\vartheta}$, where $\pi_{At} \equiv P_{At}/P_{At-1}$ is PPI inflation in country A.

Each household's real labor income (gross of taxes) is given by $w_t^p n_t^{i,p} l_t^{i,p} + w_t^g n_t^{i,g} l_t^{i,g}$, where w_t^p is the hourly real wage paid in the private sector (to be derived later) and w_t^g is the hourly real wage of the government sector. The labor income tax rate is denoted by τ_t^w. Household members who are unemployed receive unemployment benefits κ_t^B. τ_t^c denotes the consumption tax rate and T^i are constant steady-state lump-sum taxes (or, if negative, subsidies) the different household types have to pay (receive) in steady state.

Optimizing households can further invest in physical capital, domestic government bonds, or international assets. Investments in physical capital k_t^o earn a real rental rate r_t^k, while the capital depreciates at rate δ^k. Returns on physical capital net of depreciation allowances are taxed at rate τ_t^k. Nominal government bonds B_t^o pay a gross nominal interest rate R_t. Finally, D_t^o denote holdings of international nominal bonds, which pay the gross nominal interest rate R_t^{ecb}.[4] Π_t^o are nominal per capita profits generated by firms net of vacancy posting costs. We assume that all firms are owned by the

[4] In order to ensure stationarity of international bond holdings, we follow Schmitt-Grohé and Uribe (2003) and assume that there exist portfolio adjustment costs of the form $\psi_d/2\left(d_t - \bar{d}\right)^2$, with $\psi_d > 0$ and $d_t \equiv D_t/P_t$. We assume

optimizing households and that profits are redistributed in a lump-sum manner. Summarizing, the optimizers' period-budget constraint in real terms is

$$\begin{aligned}(1+\tau^c)c_t^o + I_t^o + \frac{B_t^o + D_t^o}{P_t} + T^o \\ = \frac{\Pi_t^o}{P_t} + \left((1-\tau^k)r_t^k + \tau^k\delta^k\right)k_{t-1}^o \\ + \frac{R_{t-1}B_{t-1}^o}{P_t} + \frac{R_{t-1}^{ecb}D_{t-1}^o}{P_t} - \frac{\psi_d}{2}\cdot\left(\frac{D_t^o}{P_t} - \frac{\bar{D}^o}{\bar{P}}\right)^2 \\ + (1-\tau^w)\left(w_t^p n_t^{o,p} l_t^{o,p} + w_t^g n_t^{o,g} l_t^{o,g}\right) + u_t^o \kappa_t^B,\end{aligned} \tag{2}$$

while the one for RoT consumers is given by

$$(1+\tau^c)c_t^r + T^r = (1-\tau^w)\left(w_t^p n_t^{r,p} l_t^{r,p} + w_t^g n_t^{r,g} l_t^{r,g}\right) + u_t^r \kappa_t^B. \tag{3}$$

Taking into account that RoT households do not own physical capital, the capital law of motion is given by $k_t^o = (1-\delta^k)k_{t-1}^o + \left[1 - S\left(I_t^o/I_{t-1}^o\right)\right] I_t^o$, where $S\left(I_t^o/I_{t-1}^o\right) = \frac{\kappa_I}{2}\left(I_t^o/I_{t-1}^o - 1\right)^2$ represents investment adjustment costs (see Christiano, Eichenbaum, and Evans 2005 for discussion). Maximizing (1) subject to the budget constraint and the capital law of motion yields standard first-order conditions for optimizing households.

3.1.2 Production

The retail and intermediate goods sectors of the economy are similar to Smets and Wouters (2003, 2007) or Christiano, Eichenbaum, and Evans (2005), with the exception that labor services are not hired directly from the households but from a sector of firms that produce homogenous labor services in the manner of Christoffel et al. (2009), de Walque et al. (2009), Boscá, Doménech, and Ferri (2011), or Stähler and Thomas (2012).

Final Goods Producer. There is a measure-ω continuum of firms in the final goods sector, in which firms purchase a variety of

---

for simplicity that trading in domestic government and in international bonds is not taxed.

differentiated intermediate goods and bundle these into a final good, which is sold under perfect competition. Assuming that the law of one price holds within the union, the price of the home country's final good is the same in both countries, equal to P_{At}. The problem of the representative retail firm reads $\max_{\{\tilde{y}_t(j):j\in[0,\omega]\}} P_{At}Y_t - \int_0^{\omega} P_{At}(j)\tilde{y}_t(j)dj$, where $Y_t = \left(\int_0^{\omega} \left(\frac{1}{\omega}\right)^{1/\epsilon} \tilde{y}_t(j)^{(\epsilon-1)/\epsilon}dj\right)^{\epsilon/(\epsilon-1)}$ with $\epsilon > 1$ is the retailer's production function, $\tilde{y}_t(j)$ is the retailer's demand for each differentiated input $j \in [0,\omega]$, and $P_{At}(j)$ is the nominal price of each input. The standard first-order condition for the problem is given by $\tilde{y}_t(j) = (P_{At}(j)/P_{At})^{-\epsilon} \frac{Y_t}{\omega}$. Combining the latter with the retailer's production function and the zero-profit condition, we obtain that the producer price index in the home country must equal $P_{At} = \left(\int_0^{\omega} \frac{1}{\omega} P_{At}(j)^{1-\epsilon}dj\right)^{1/(1-\epsilon)}$. Total demand for each intermediate input equals $\omega\tilde{y}_t(j) \equiv y_t(j) = \left(\frac{P_{At}(j)}{P_{At}}\right)^{-\epsilon} Y_t$, as there are ω retail firms.

Intermediate Goods. Each intermediate goods producer $j \in [0,\omega]$ faces the technology

$$y_t(j) = \epsilon^a \cdot \left[\tilde{k}_t(j)\right]^{\alpha} \cdot \left[\tilde{lab}_t(j)\right]^{(1-\alpha)}, \tag{4}$$

where $\alpha \in [0,1]$ is the elasticity of output with respect to capital, $\tilde{lab}_t(j)$ denotes the demand for effective labor services, $\tilde{k}_t(j)$ is the demand for effective capital, and ϵ^a is total factor productivity. Following Coenen, Straub, and Trabandt (2013), we assume that effective capital is a constant elasticity of substitution (CES) composite given by

$$\tilde{k}_t(j) = \left(\alpha_k^{\frac{1}{v_k}} \left(k_{t-1}^p(j)\right)^{\frac{v_k-1}{v_k}} + (1-\alpha_k)^{\frac{1}{v_k}} \left(k_{t-1}^g\right)^{\frac{v_k-1}{v_k}}\right)^{\frac{v_k}{v_k-1}},$$

where k_{t-1}^g is the public capital stock available in period t, which is determined by government investment. It is assumed to be productivity enhancing, where $\alpha_k \in (0,1]$ is a share parameter, and the parameter v_k denotes the elasticity of substitution between private capital services and the public capital stock (see also Pappa 2009 and Leeper, Walker, and Yang 2010 for discussion). An analogous aggregator is given for public employment,

$$\tilde{lab}_t(j) = \left(\alpha_g^{\frac{1}{v_g}} \left(N_t^p(j)\, h_t^p(j)\right)^{\frac{v_g-1}{v_g}} + (1-\alpha_g)^{\frac{1}{v_g}} \left(N_t^g\, h_t^g\right)^{\frac{v_g-1}{v_g}}\right)^{\frac{v_g}{v_g-1}},$$

following Fernández-de-Cordoba, Pérez, and Torres (2012).

Intermediate goods firms acquire private labor and capital services in perfectly competitive factor markets at real (CPI-deflated) prices x_t and r_t^k, respectively. Cost minimization subject to (4) implies the factor demand conditions for capital and labor $r_t^k = mc_t \cdot \alpha \cdot y_t(j)/\tilde{k}_t(j) \cdot \partial\tilde{k}_t(j)/\partial k_{t-1}^p(j)$ and $x_t = mc_t \cdot (1-\alpha) \cdot y_t(j)/\tilde{lab}_t(j) \cdot \partial\tilde{lab}_t(j)/\partial N_t^p(j)$, where mc_t is the real (CPI-deflated) marginal cost common to all intermediate good producers. The capital-labor ratios are equalized across firms because of constant returns to scale in capital and labor and perfectly competitive (private) input prices.

As is standard in the literature, intermediate goods firms set nominal prices à la Calvo (1983). This implies that a randomly chosen fraction $\theta_P \in [0,1)$ of firms cannot reoptimize their price in each period. A firm that has the chance to reoptimize its price in period t maximizes $E_t \sum_{z=0}^{\infty} (\beta\theta_P)^z \frac{\lambda_{t+z}^o}{\lambda_t^o} \left[\frac{P_{At}(j)}{P_{t+z}} - mc_{t+z}\right] y_{t+z}(j)$ with respect to the nominal price $P_{At}(j)$, subject to $y_{t+z}(j) = (P_{At}(j)/P_{At+z})^{-\epsilon}\, Y_{t+z}$. λ_t^i represents the marginal consumption utility of households of type i. The first-order condition is standard and implies the standard law of motion for the price level, $1 = \theta_P \left(\frac{1}{\pi_{At}}\right)^{1-\epsilon} + (1-\theta_P)\tilde{p}_t^{1-\epsilon}$, where $\tilde{p}_t \equiv \tilde{P}_{At}/P_{At}$ is the relative (PPI-deflated) optimal price and $\tilde{P}_{At}$ is the optimal price chosen by all period-t price setters.

3.1.3 The Labor Market

Following Christoffel et al. (2009), de Walque et al. (2009), and Stähler and Thomas (2012), we assume that labor firms hire workers from the household sector in order to produce homogenous labor services, which they sell to intermediate goods producers at the perfectly competitive price x_t. The production function of each labor firm is linear in the number of hours worked by its employee. With

N_t^p being the fraction of the total labor force employed in the private sector and the fact that optimizers and RoTs will work the same amount of hours (which we show below), the total per capita supply of labor services is given by $Lab_t = N_t^P \cdot l_t^p$. Equilibrium in the market for labor services requires that $\omega Lab_t = \int_0^\omega lab_t(j)dj$.

Using demand for each intermediate input and the production function (4) plus the fact that the capital-labor ratio is equalized across intermediate goods firms, this yields $Y_t D_t = \epsilon^a \tilde{k}_t^\alpha \tilde{Lab}_t^{1-\alpha}$, where $D_t \equiv \int_0^\omega \omega^{-1} \left(P_{At}(j)/P_{At}\right)^{-\epsilon} dj$ is a measure of price dispersion. In what follows, we will specify the matching process, flows in the labor market, private-sector vacancy creation, the corresponding wage determination, and labor market participation decisions. Government wages and employment are autonomously chosen by the fiscal authority (see section 3.1.4).

Matching Process and Labor Market Flows. A household member can be in one of three states: (i) employed in the public sector, (ii) employed in the private sector, or (iii) unemployed. Unemployment is the residual state in the sense that a worker whose employment relationship ends flows into unemployment. All unemployed workers search for a job. We assume that searchers are randomly matched to the private or the public sector.

Denoting total sector-specific per capita employment in period t by $N_t^f = (1-\mu)n_t^{o,f} + \mu n_t^{r,f}$, where $f = p, g$ stands for private and government employment, the total economy-wide employment rate is given by $N_t^{tot} = N_t^p + N_t^g$, and the aggregate unemployment rate is given by $U_t = 1 - N_t^{tot}$. Following Blanchard and Galí (2010), we assume that the hiring round takes place at the beginning of each period, and that new hires start producing immediately. We also assume that workers, who are dismissed at the end of period $t-1$, start searching for a new job at the beginning of period t. Therefore, the pool of searching workers at the beginning of period t is given by

$$\tilde{U}_t = U_{t-1} + s^p N_{t-1}^p + s^g N_{t-1}^g,$$

where s^f, with $f = p, g$, represents the constant separation rate in the private (p) and public (g) sector.

The matching process is governed by a standard Cobb-Douglas aggregate matching function for each sector $f = p, g$, $M_t^f = \kappa_e^f \cdot$

$\left(\tilde{U}_t\right)^{\varphi^f} \cdot \left(v_t^f\right)^{(1-\varphi^f)}$, where $\kappa_e^f > 0$ is the sector-specific matching efficiency parameter, $\varphi^f \in (0,1)$ the sector-specific matching elasticity, and M_t^f the number of new matches formed in period t resulting from the total number of searchers and the number of sector-specific vacancies v_t^f. The probability of an unemployed worker finding a job in sector f can thus be stated as $p_t^f = M_t^f/\tilde{U}_t$, while the probability of filling a vacancy is given by $q_t^f = M_t^f/v_t^f$. With the constant separation rate in each sector, the law of motion for sector-specific employment rates is therefore given by

$$N_t^f = \left(1 - s^f\right) \cdot N_{t-1}^f + p_t^f \cdot \tilde{U}_t. \tag{5}$$

Thus, employment in sector f today is given by yesterday's employment that has not been destroyed plus newly created matches in that sector.

Asset Values of Jobs, Wage Bargaining, and Job Creation. As is standard in the literature, we assume that firms and workers bargain about their share of the overall match surplus to determine wages and hours. Following Boscá, Doménech, and Ferri (2009, 2011) and Boscá et al. (2010), we assume that a union, which takes into account (aggregate) utility of optimizing and RoT households, undertakes the bargaining. Furthermore, we assume staggered bargaining of nominal wages similar to Gertler, Sala, and Trigari (2008). This implies that, each period, a randomly chosen fraction θ_w of continuing firms cannot renegotiate wages and hours, while a fraction θ_w^n of newly created firms does not bargain either and is stuck having to pay the previous period's average nominal wage for the average hours worked of the previous period. When letting $J_t\left(\tilde{W}_t^p\right)$ be the value function of employment for firms that are allowed to bargain and $\Omega_t \equiv (1-\mu)H_t^{o,p}\left(\tilde{W}_t^p\right) + \mu H_t^{r,p}\left(\tilde{W}_t^p\right)$ that of the union, where $H_t^{i,p}\left(\tilde{W}_t^p\right)$ is the corresponding household type-i utility, the Nash problem is given by

$$\max_{\tilde{W}_t^p, \tilde{l}_t^p} \; [\Omega_t]^{\xi} \left[J_t\left(\tilde{W}_t^p\right)\right]^{1-\xi}, \tag{6}$$

where $\xi \in [0,1)$ is the union's bargaining power, $\tilde{W}_t^p$ denotes the nominal wage negotiated in period t, and $\tilde{l}_t^p$ denotes the corresponding amount of hours worked. The value function of a firm that renegotiates in that period is given by

$$J_t\left(\tilde{W}_t^p\right) = E_t \sum_{k=0}^{\infty} \left\{ [\beta \cdot (1-s^p) \cdot \theta_w]^k \cdot \frac{\lambda_{t+k}^o}{\lambda_t^o} \right.$$
$$\left. \cdot \left[x_{t+k} - (1+\tau_{t+k}^{sc}) \cdot \frac{\tilde{W}_t^p}{P_{t+k}} \right] \cdot \tilde{l}_t^p \right\} + (1-\theta_w)$$
$$\cdot E_t \sum_{k=1}^{\infty} \left\{ [\beta \cdot (1-s^p)]^k \cdot \theta_w^{k-1} \cdot \frac{\lambda_{t+k}^o}{\lambda_t^o} \cdot J_{t+k}\left(\tilde{W}_{t+k}^p\right) \right\},$$

where τ_t^{sc} is the social security contribution rate. The value of the firm is the discounted profit flow in those future states in which it is not allowed to renegotiate plus its continuation value should it have the chance to reoptimize in the next period. For new jobs where firm and worker do not bargain, the nominal wage equals last period's average nominal wage, W_{t-1}^p, the amount of hours is given by l_{t-1}^p, and the value of the job equals

$$J_t\left(W_{t-1}^p\right) = J_t\left(\tilde{W}_t^p\right) - E_t \sum_{k=0}^{\infty} \left\{ [\beta \cdot (1-s^p) \cdot \theta_w]^k \cdot \frac{\lambda_{t+k}^o}{\lambda_t^o} \right.$$
$$\left. \cdot (1+\tau_{t+k}^{sc}) \cdot \frac{W_{t-1}^p \cdot l_{t-1}^p - \tilde{W}_t^p \cdot \tilde{l}_t^p}{P_{t+k}} \right\}.$$

Analogously, we can derive how workers value a match surplus. Since different household types use different stochastic discount factors,[5]

[5]Actually, RoT consumers are not allowed to use their wealth to smooth consumption over time. But they can take advantage of the fact that a matching today may continue in the future, yielding a labor income that affects consumption tomorrow. Therefore, RoT households use the margin that hours and wage negotiations provide them to improve their lifetime utility by narrowing the gap in utility with respect to optimizers. In this sense, they compare their intertemporal marginal rate of substitution (their "stochastic discount factor") with the one of optimizers; see Boscá, Doménech, and Ferri (2011) for a more detailed discussion.

we must distinguish between the surplus for an optimizing and a rule-of-thumb household. For a worker belonging to a type-i household, the surplus value of a job in a renegotiating firm is given by

$$H_t^{i,p}\left(\tilde{W}_t^p\right) = E_t \sum_{k=0}^{\infty} \left\{ [\beta \cdot (1-s^p) \cdot \theta_w]^k \cdot \frac{\lambda_{t+k}^i}{\lambda_t^i} \cdot \left[(1-\tau_{t+k}^w) \cdot \frac{\tilde{W}_t^p}{P_{t+k}} \right.\right.$$
$$\left.\left. \cdot\, \tilde{l}_t^p - \kappa^h \cdot \frac{\tilde{l}_t^{p\ 1+\sigma_h}}{(1+\sigma_h)\lambda_{t+k}^i} - \Xi_{t+k}^{i,p} \right] \right\} + (1-\theta_w)$$
$$\cdot\, E_t \sum_{k=1}^{\infty} \left\{ [\beta \cdot (1-s^p)]^k \cdot \theta_w^{k-1} \cdot \frac{\lambda_{t+k}^i}{\lambda_t^i} \cdot H_{t+k}^{i,p}(\tilde{W}_{t+k}^p) \right\},$$

for $i = o, r$, where

$$\Xi_t^{i,f} = \kappa_t^B + \beta(1-s^f)E_t \frac{\lambda_{t+1}^i}{\lambda_t^i}$$
$$\cdot \left\{ p_{t+1}^g H_{t+1}^{i,g} + p_{t+1}^p \left[(1-\theta_w^n) H_{t+1}^{i,p}\left(\tilde{W}_{t+1}^p\right) + \theta_w^n H_{t+1}^{i,p}\left(W_t^p\right) \right] \right\}$$

represents the outside option of a type-i worker employed in sector $f = p, g$ at time t. The latter is the sum of unemployment benefits, κ_t^B, and the expected value of searching for a job in the following period, where p_{t+1}^f is the probability of finding a job in sector $f = p, g$. Conditional on landing on a private-sector job ($f = p$), the surplus value for the worker is contingent on whether the firm is allowed to bargain (in which case the worker receives $\tilde{W}_{t+1}^p$, and works $\tilde{l}_{t+1}^p$ hours) or not (in which case she receives today's average wage, W_t^p, and works l_t^p hours). In new jobs where the wage and hours are not optimally bargained, the surplus value enjoyed by type-i workers is given by

$$H_t^{i,p}\left(W_{t-1}^p\right) = H_t^{i,p}\left(\tilde{W}_t^p\right) + E_t \sum_{k=0}^{\infty} \left\{ [\beta \cdot (1-s^p) \cdot \theta_w]^k \right.$$
$$\left. \cdot\, \frac{\lambda_{t+k}^i}{\lambda_t^i} \cdot (1-\tau_{t+k}^w) \cdot \frac{W_{t-1}^p \cdot l_{t-1}^p - \tilde{W}_t^p \cdot \tilde{l}_t^p}{P_{t+k}} \right\}.$$

Note that $H_t^{i,g}$ denotes the surplus value of a government job for a type-i worker. As wages and hours there are autonomously set by the fiscal authority, the asset value function simplifies to

$$H_t^{i,g} = (1-\tau_t^w)w_t^g \cdot l_t^g - \Xi_t^{i,g} - \kappa^h \cdot \frac{l_t^{g\,1+\sigma_h}}{1+\sigma_h} + \beta(1-s^g)E_t\left\{\frac{\lambda_{t+1}^i}{\lambda_t^i}\cdot H_{t+1}^{i,g}\right\},$$

where w_t^g is the real wage paid by the government and l_t^g the amount of hours a worker employed by the government has to work. Given the asset value functions of firms and workers, we are now in a position to solve the wage-bargaining game (6). The resulting sharing rule is given by

$$\Omega_t = \frac{\xi}{1-\xi} \cdot \frac{E_t\sum_{z=0}^{\infty}\left\{\left((1-\mu)\frac{\lambda_{t+z}^o}{\lambda_t^o}+\mu\frac{\lambda_{t+z}^r}{\lambda_t^r}\right)[\beta(1-s^p)\theta_w]^z\,\frac{(1-\tau_{t+z}^w)}{P_{t+z}}\right\}}{E_t\sum_{z=0}^{\infty}\left\{\frac{\lambda_{t+z}^o}{\lambda_t^o}[\beta(1-s^p)\theta_w]^z\,\frac{(1+\tau_{t+z}^{sc})}{P_{t+z}}\right\}} \cdot J_t\left(\tilde{W}_t^p\right). \tag{7}$$

Solving equation (7) for $\tilde{W}_t^p$ by using the corresponding asset value functions gives the optimal wage bargained in period t. The average real wage in the private sector, $w_t^p \equiv W_t^p/P_t$, hence evolves according to

$$w_t^p = \frac{(1-s^p)N_{t-1}^p}{N_t^p}\left[(1-\theta_w)\tilde{w}_t^p + \theta_w\cdot\frac{w_{t-1}^p}{\pi_t}\right] + \frac{M_t^P}{N_t^p}\left[(1-\theta_w^n)\tilde{w}_t^p + \theta_w^n\cdot\frac{w_{t-1}^p}{\pi_t}\right], \tag{8}$$

where $\tilde{w}_t^p \equiv \tilde{W}_t^p/P_t$ is the real optimally bargained wage and $w_{t-1}^p/\pi_t = W_{t-1}^p/P_t$ is the real value of yesterday's average nominal wage at today's prices. We have also taken into account the fact that new and continuing jobs pay the optimally bargained wage with probabilities $1-\theta_w^n$ and $1-\theta_w$, respectively.

For the hours determination in the private sector, we get

$$AA_t \cdot \frac{E_t \sum_{z=0}^{\infty} \left\{ \left((1-\mu) \frac{\lambda^o_{t+z}}{\lambda^o_t} + \mu \frac{\lambda^r_{t+z}}{\lambda^r_t} \right) \left[\beta (1-s^p) \theta_w \right]^z \frac{(1-\tau^w_{t+z})}{P_{t+z}} \right\}}{E_t \sum_{z=0}^{\infty} \left\{ \frac{\lambda^o_{t+z}}{\lambda^o_t} \left[\beta (1-s^p) \theta_w \right]^z \frac{(1+\tau^{sc}_{t+z})}{P_{t+z}} \right\}}$$
$$= E_t \sum_{z=0}^{\infty} \left\{ \left((1-\mu) \frac{\kappa^h}{\lambda^o_t} + \mu \frac{\kappa^h}{\lambda^r_t} \right) \left[\beta (1-s^p) \theta_w \right]^z \right\} \cdot \tilde{l}^{p \ \sigma_h}_t , \tag{9}$$

where $AA_t = E_t \sum_{z=0}^{\infty} \left\{ \frac{\lambda^o_{t+z}}{\lambda^o_t} \left[\beta (1-s^p) \theta_w \right]^z x_{t+z} \right\}$. Average hours worked in period t are analogously aggregated as wages; see equation (8).

It remains to determine how jobs are created. As is standard in the literature, we assume that opening a vacancy has a real (CPI-deflated) flow cost of κ^p_v. Following Pissarides (2009), we further assume that free entry into the vacancy posting market drives the expected value of a vacancy to zero. Under our assumption of instantaneous hiring, real vacancy posting costs, κ^p_v, must equal the time-t vacancy filling probability, q^p_t, times the expected value of a filled job in period t net of training costs. The latter condition can be expressed as

$$\frac{\kappa^p_v}{q^p_t} = (1-\theta^n_w) \cdot J_t \left(\tilde{W}^p_t \right) + \theta^n_w \cdot J_t \left(W^p_{t-1} \right), \tag{10}$$

where we take into account that the wage of the newly created job may be optimally bargained with probability $1-\theta^n_w$.

3.1.4 Fiscal Authorities

Defining the (CPI-deflated) per capita value of end-of-period government debt as $b_t \equiv B_t/P_t$, we can state that it evolves according to a standard debt accumulation equation, $b_t = \frac{R_{t-1}}{\pi_t} b_{t-1} + PD_t$, where PD_t denotes real (CPI-deflated) per capita primary deficit. The latter is given by per capita fiscal expenditures minus per capita fiscal revenues,

$$
\begin{aligned}
PD_t = & \left[\frac{G_t}{p_{Bt}^{1-\omega-\psi}} + \kappa_t^B U_t + \kappa_v^g v_t^g + Sub_t \right] \\
& - \left[(\tau_t^w + \tau_t^{sc}) \left[w_t^p N_t^p l_t^p + w_t^g N_t^g l_t^g \right] \right. \\
& \left. + \tau_t^c C_t + \tau_t^k (r_t^k - \delta^k) k_{t-1} + (1-\mu) T^o + \mu T^r \right],
\end{aligned}
$$

where G_t denotes per capita government spending in investment, consumption goods, and services expressed in PPI terms (hence the correction for the CPI-to-PPI ratio, $P_t/P_{At} = p_{Bt}^{1-\omega-\psi}$). Letting C_t^g and I_t^g denote real per capita public purchases and public investment, respectively, we have the following nominal relationship: $P_{At} G_t = P_{At} (C_t^g + I_t^g) + (1 + \tau_t^{sc}) P_t w_t^g N_t^g l_t^g$. Dividing by P_{At} and using $P_t/P_{At} = p_{Bt}^{1-\omega-\psi}$, we obtain $G_t = C_t^g + I_t^g + [(1 + \tau_t^{sc}) w_t^g N_t^g l_t^g] \, p_{Bt}^{1-\omega-\psi}$.

We assume that $\kappa_t^B = rrs \cdot (1 - \bar{\tau}^w) \bar{w}^p \bar{l}^p$. Here, rrs is then the unemployment benefit replacement ratio and the bar indicates (initial) steady-state values. Given public investment, the stock of public physical capital evolves as follows: $k_t^g = (1 - \delta^g) k_{t-1}^g + I_t^g$, where we assume that the public capital stock depreciates at rate δ^g. To guarantee stationarity of public debt, for *at least* one fiscal instrument $X \in \{\tau^w, \tau^{sc}, \tau^c, C^g, N^g\}$, the government must follow a fiscal rule of the form

$$
X_t = \bar{X} + \rho_X \left(X_{t-1} - \bar{X} \right) + (1 - \rho_X) \, \phi_X \cdot \left(\frac{b_{t-1}}{Y_{t-1}^{tot}} p_{Bt-1}^{1-\omega-\psi} - \omega^b \right) + \epsilon_t^X, \tag{11}
$$

in which the coefficient ϕ_X, i.e., fiscal policy's stance on debt deviations from target, is non-zero (positive for revenue instruments, negative for expenditure instruments). ρ_X is a smoothing parameter. Following Galí, Lopez-Salido, and Vallés (2007), T^r is chosen such that, in steady state, optimizing and RoT households consume the same, while T^o closes the government's budget.

3.1.5 International Linkages and Union-Wide Monetary Policy

This section describes the international linkages via trade in goods and foreign assets, market clearing, and the union-wide monetary policy rule.

International Linkages. International linkages between the two countries are given by trade in goods and services as well as in international bonds. The home country's net foreign asset position, expressed in terms of PPI, evolves according to

$$d_t = \frac{R_{t-1}^{ecb} \cdot d_{t-1}}{\pi_{At}} + \frac{1-\omega}{\omega}\left(C_{At}^{*} + I_{At}^{*}\right) - p_{Bt}\left(C_{Bt} + I_{Bt}\right), \qquad (12)$$

where $(1-\omega)\left(C_{At}^{*} + I_{At}^{*}\right)/\omega$ are real per capita exports and $p_{Bt}\left(C_{Bt} + I_{Bt}\right)$ are real per capita imports. Zero net supply of international bonds implies $\omega d_t + (1-\omega)\, p_t^B d_t^* = 0$. Terms of trade $p_{Bt} = P_{Bt}/P_{At}$ evolve according to $p_{Bt} = (\pi_{Bt}/\pi_{At})\, p_{Bt-1}$.

Equilibrium in Goods Markets and GDP. Market clearing implies that private per capita production in the home and foreign country, Y_t and Y_t^*, respectively, is used for private and public consumption and private and public investment demand as well as private and public vacancy posting costs,

$$\begin{aligned} Y_t &= C_{At} + I_{At} + C_t^g + I_t^g + \frac{1-\omega}{\omega}\left(C_{At}^{*} + I_{At}^{*}\right) \\ &\quad + {p_t^B}^{\,1-\omega-\Psi}\, \kappa_v^p \left(v_t^p + v_t^g\right), \qquad (13) \\ Y_t^* &= C_{Bt}^{*} + I_{Bt}^{*} + C_t^{g*} + I_t^{g*} + \frac{\omega}{1-\omega}\left(C_{Bt} + I_{Bt}\right) \\ &\quad + \left(1/p_t^B\right)^{\omega-\Psi^*} \kappa_v^{p*} \left(v_t^{p*} + v_t^{g*}\right), \qquad (14) \end{aligned}$$

where we have assumed that vacancy posting costs in the private and public sector are the same, $\kappa_v^g = \kappa_v^p$. Consistent with national accounting and in line with Stähler and Thomas (2012), each country's GDP is the sum of private-sector production and government production of goods and services. The latter is measured at input costs, that is, by the gross government wage bill. Hence, home and foreign real (PPI-deflated) per capita GDP are given by $Y_t^{tot} = Y_t + (1+\tau_t^{sc}) w_t^g N_t^g l_t^g p_{Bt}^{1-\omega-\psi}$ and $Y_t^{tot,*} = Y_t^* + (1+\tau_t^{sc*}) w_t^{g*} N_t^{g*} l_t^{g*} p_{Bt}^{-(\omega-\psi^*)}$, respectively.

Monetary Authority. We assume that the area-wide monetary authority has its nominal interest rate, R_t^{ecb}, respond to deviations of area-wide inflation from its long-run target, $\bar{\pi}$, and to area-wide GDP growth, according to a simple Taylor rule,

$$\frac{R_t^{ecb}}{\bar{R}^{ecb}} = \left(\frac{R_{t-1}^{ecb}}{\bar{R}^{ecb}}\right)^{\rho_R} \cdot \left\{ \left[\left(\frac{\pi_t}{\bar{\pi}}\right)^{\omega} \left(\frac{\pi_t^*}{\bar{\pi}^*}\right)^{1-\omega} \right]^{\phi_\pi} \left[\left(\frac{Y_t}{\bar{Y}}\right)^{\omega} \left(\frac{Y_t^*}{\bar{Y}*}\right)^{1-\omega} \right]^{\phi_y} \right\}^{(1-\rho_R)},$$

where ρ_R is a smoothing parameter, and ϕ_π and ϕ_y are the monetary policy's stance on inflation and output growth, respectively.

3.1.6 Welfare

In order to assess welfare effects of the reform measures, we compute the lifetime consumption-equivalent gain of each type of household as a result of the change in fiscal policy.[6] We will take into account the welfare difference between the initial and the final steady state as well as the transition thereto. More precisely, we calculate the consumption-equivalent welfare gain, ce^i, such that

$$\sum_{t=0}^{\infty} \left(\beta^i\right)^t U\left((1+ce^i)\bar{c}^i, \bar{l}^p, \bar{l}^g\right) = \sum_{t=0}^{\infty} \left(\beta^i\right)^t U\left(c_t^i, l_t^p, l_t^g\right),$$

where the utility function $U(\cdot)$ is given by equation (1) and the bar indicates initial steady-state values. Hence, ce^i represents the amount of initial steady-state consumption a household of type i is willing to give up in order to live in the alternative regime after the policy change. Economy-wide welfare is computed as $ce^{tot} = (1-\mu)ce^o + \mu ce^r$.

This completes the model description. We now turn to the model calibration.

[6] Among the large literature using consumption equivalents for welfare comparison, see, for example, Obstfeld (1994), Otrok (2001), Krebs (2003), Lucas (2003), Barro (2006), and Cristoffel et al. (2009).

3.2 Calibration

The model is calibrated to quarterly frequency. The home country represents EMU's periphery and the foreign country its core. In what follows, we use the term "country" in the model sense and use the words "home"/"foreign" and "periphery"/"core" interchangeably.[7]

For the general calibration strategy, we broadly rely on Stähler and Thomas (2012). This means that our strategy consists of (i) matching some steady-state variables with their counterparts in the data and (ii) carefully choosing the remaining free parameter values in line with the existing literature. The data we use are largely based on a data set ranging from 1999:Q1 to 2013:Q4 for the euro area containing a rich set of quarterly fiscal variables, described in more detail in Gadatsch, Hauzenberger, and Stähler (2016). The primary sources for the various variables are the European System of Accounts (ESA) for the main aggregates and the European Commission for the fiscal variables. Some labor market variables come from OECD data. Hence, in the initial steady state, we match data averages with the corresponding model variables. Furthermore, given an import share of 15 percent in the periphery (see Balta and Delgado 2009), we normalize periphery per capita GDP, PPI inflation, and the terms of trade to one and set the net foreign asset position to zero. Then, we target home-country import and export shares vis-à-vis the euro area, which is the reason we have to derive the corresponding home bias parameters endogenously. We also set the foreign-country per capita GDP relative to the home-country per-capita GDP according to the relative share of total GDP in EMU. Vacancy filling rates for the euro area are estimated in Christoffel et al. (2009) and assumed to be equal across countries due to the lack of reliable data. The calibration of the labor market parameters is also similar to Moyen, Stähler, and Winkler (2019). Furthermore, we normalize total time available to a household member to one and, hence, assume that, once employed, one-third of total time is devoted to work in the initial steady state, which is a standard

[7]Following Moyen, Stähler, and Winkler (2019), we calibrate the model to seven euro-zone core countries (Austria, Belgium, Germany, Finland, France, Luxembourg, Netherlands) and five periphery countries (Spain, Greece, Ireland, Italy, Portugal).

Table 1. Targeted Steady-State Variables in Baseline Calibration

Targeted Variable	Periphery	Core
Relative Population	0.399	0.601
Real GDP Share	0.347	0.653
Imports-to-GDP Ratio	0.150	NA
Share of Liquidity-Constrained Consumers	0.500	0.460
Public Revenues		
Consumption Tax Rate	0.196	0.183
Capital Tax Rate	0.316	0.214
Labor Income Tax Rate	0.277	0.304
SSC Rate	0.246	0.167
Public Spending		
Gov. Purchases-to-GDP Ratio	0.120	0.110
Gov. Investment-to-GDP Ratio	0.020	0.020
Public-Sector Wage Bill	0.080	0.080
Debt-to-GDP Ratio	0.900	0.750
Labor Market		
Unemployment Rate	0.122	0.084
Public Employees/Total Employment	0.180	0.160
Premium of Public over Private Wages	0.080	0.060
Replacement Rate (Unemployment Benefits)	0.523	0.690
Vacancy Filling Rate (Private)	0.700	0.700
Vacancy Filling Rate (Public)	0.800	0.800

Notes: Tax rates are implicit tax rates. The “NA” for the imports-to-GDP ratio in the rest of the euro area (RoE) is due to the fact that this needs to be derived “endogenously” to match a steady-state net foreign asset position of zero together with a steady-state real exchange rate of one (see table 2).

assumption in the literature. The target values are summarized in table 1.

In choosing the general model parameters, we strongly rely on Christoffel et al. (2009), who estimate a model with a search-and-matching labor market to European data. Note that the simulation results are highly robust to alternative parameter calibration. The discount factor is set to $\beta = 0.992$ to match an annual real rate of 3.2 percent. Risk aversion $\sigma_c = 2$ as well as habits in consumption $h = 0.6$ are set close to the mode estimates in Smets and

Wouters (2003). The share of RoTs, μ, is set to 0.5 in the periphery and 0.46 in the core following Le Blanc et al. (2014). Forni, Monteforte, and Sessa (2009) find similar values for the overall euro area.

We assume that, in the initial steady state, optimizers and RoT households face the same consumption level as in Bilbiie (2008), which determines the values for T^o and T^r. Monetary policy parameters are standard values of a conventional Taylor rule, while the price markup and the Calvo parameters for prices and wages are set in line with estimates from the New Area-Wide Model (see Christoffel, Coenen, and Warne 2008 for a discussion). Capital depreciation is set to a standard value of $\delta^p = \delta^g = 0.025$ and the capital share in production is set to one-third (Cooley and Prescott 1995), while capital adjustment costs are set to a standard value close to 5. For the CES aggregator of private and public capital, we rely on the estimates of Coenen, Straub, and Trabandt (2013), i.e., we set $\alpha_k = 0.9$ and $v_k = 0.84$. Similar values are chosen for the CES aggregator of private and public employment. According to Schmitt-Grohé and Uribe (2003), it is sufficient to choose a rather small value for the risk premium parameter on international bonds in order to generate a stable equilibrium. So we opt for $\Psi_d = \Psi_d^* = 0.01$.

Regarding the labor market, the elasticity of the matching function in the private sector, φ^p, is set to 0.5 in line with Petrongolo and Pissarides (2001), Burda and Weder (2002), and Christoffel et al. (2009). The value in the public sector, φ^g, is set a bit lower, to 0.3, following Afonso and Gomes (2014). The bargaining power of workers is derived endogenously to match the premium of public over private wages. The quarterly separation rate in the private sector is set to 0.04 in line with Christoffel et al. (2009). Again, it is a bit lower in the public sector. For nominal wage rigidities, Colciago et al. (2008), Christoffel et al. (2009), and de Walque et al. (2009) find a rather high degree of stickiness (note that the latter paper is based on U.S. data, however). We opt for a middle value of these studies and set $\theta_w = \theta_w^n = 0.83$.

Given these parameters, it remains to derive the efficiency of the matching function as well as vacancy posting costs endogenously to meet the targeted labor market variables shown in table 1. Values derived endogenously to match the targets are marked by the superscript e in table 2. We perform robustness analysis regarding trade

Table 2. Baseline Parameter Calibration

Parameter	**Symbol**	**Value**
Monetary Policy		
Interest Rate Smoothing	ρ_R	0.850
Stance on Inflation	ϕ_π	1.500
Stance on Output Gap	ϕ_y	0.125
Fiscal Policy		
Fiscal Smoothing Parameter	ρ_X	0.900
Stance on Debt	ϕ_X	0.050
Price and Wage Stickiness		
Calvo Parameter (Prices)	θ_P	0.750
Market Power (Markup)	ϵ	4.000
Calvo Parameter (Existing Wages)	θ_w	0.830
Calvo Parameter (New Wages)	θ_w^n	0.830
Preferences		
Discount Rate	β	0.992
Risk Aversion	σ_c	2.000
Habits in Consumption	h	0.600
Trade in International Bonds		
Risk Premium Parameter	$\psi_d = \psi_d^*$	0.010
Production		
Private-Sector Capital Depreciation	δ^k	0.025
Public-Sector Capital Depreciation	δ^g	0.025
Private-Sector Capital Share in Prod.	α	0.330
Public-Sector Capital/Employment Influence in Private Production	α_k, α_g	0.900
Substitutability Public/Private Capital/Employment	υ_k, υ_g	0.840
Adjustment Cost Parameter	κ_I	4.940
Labor Market		
Matching Elasticity (Private Sector)	φ^p	0.500
Matching Elasticity (Public Sector)	φ^g	0.300
Separation Rate (Public Sector)	s^g	0.020
Separation Rate (Private Sector)	s^p	0.040
Labor Market Matching Efficiency[e]	κ_e^p	0.360
Vacancy Posting Costs[e]	κ_v^p	2.040
Bargaining Power of Workers[e]	ξ	0.200

openness, labor market efficiency, and country size. As can be seen in the results section, our results are robust to parameter changes.

4. Simulation Design

In order to assess the macroeconomic effects of reducing the tax wedge, we calibrate a reduction in the labor income tax and/or social security contribution rates that yields an increase in the governments' primary deficit-to-GDP ratio by 1 percentage point ex ante, that is, holding constant everything other than changes in the stated instruments. A higher (lower) reduction in the tax/contribution rates would, naturally, imply stronger (weaker) long-run effects. In order to compensate for the revenue losses, we then calculate an increase in other revenue components (consumption taxes) or a decrease in the expenditure components (public purchases or employment) such that the budget will be balanced ex post in the new steady state which takes into account the "second-round effects" resulting from changes in endogenous variables.[8]

For all simulations, we assume that the economy is initially in steady state. Each fiscal measure is implemented by changing the corresponding long-run target (parameter) such that the measure is permanent. We then derive the final steady state that arises after the policy change and calculate the transition from the initial to the final steady state under perfect foresight. In this calculation, we assume that no other shocks hit the economy during the transition path, which allows us to attribute all the effects to the policy measure.

[8]Technically, this is done by allowing only the steady-state value of the financing fiscal instrument that we consider to adjust according to equation (11). Assuming the targeted debt-to-GDP ratio to remain unchanged, this calculates the final steady-state value of the financing instrument which balances the government's budget and takes into account the "second-round effects." If we assume that the public-sector budget constraint is balanced ex ante (i.e., ignore the second-round effects and assume lump-sum taxes to stabilize debt in the long run), the necessary increase in the consumption tax rate must be larger and the positive effect turn out to be smaller. While the results are, overall, equivalent, this hurts RoT consumers along the transition after a cut in social security contributions because the resulting wage increase is not sufficient to immediately compensate for the policy-induced rise in consumption costs. Results of such an analysis can be found in an earlier version of the paper (see Attinasi et al. 2016).

5. Simulation Results

In this section, we present long-run effects and transition dynamics for each reform measure described above. Results are reported in percentage deviations of key macroeconomic variables from their initial steady-state values (percentage-point deviations for rates and ratios). To get a better understanding of the transmission mechanisms, we will show impulse response functions (IRFs) for selected simulations.

5.1 Reducing the Tax Wedge and Different Financing Schemes

In a first step, we assess the effects of a reduction in the labor tax wedge implemented as a reduction in either the personal income tax (PIT) paid by employees or the social security contributions (SSC) paid by the employers, assuming that it is financed by an increase in the consumption tax rate. In a second step we look at the implications of using expenditure instruments to finance the corresponding revenue loss.

5.1.1 Fiscal Devaluation: Comparing Workers' and Firms' Tax Rate Reductions

Table 3 displays the long-run changes of selected key macroeconomic variables after a fiscal devaluation episode in the periphery when (i) reducing the workers' personal labor income tax rate and when (ii) reducing the firms' social security contribution rate only. In these simulations, it is assumed that the reduction in the labor tax wedge is always financed by an appropriate increase in the consumption tax rate. The transition dynamics of selected key macroeconomic variables are shown in figure 1. Furthermore, table 3 shows the results of simulations with a lower home bias, a lower labor market efficiency, or a smaller country size.[9]

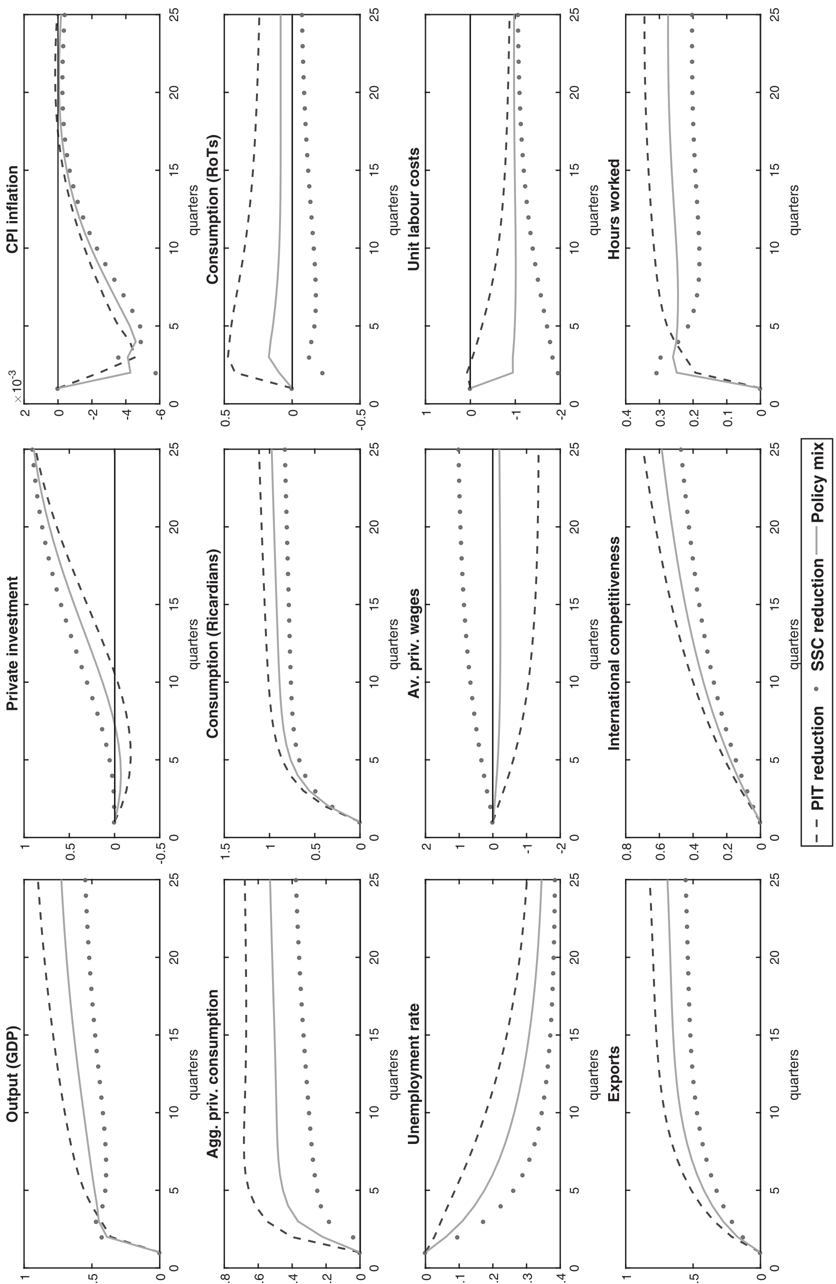
[9] For the robustness analyses, we assume that the home bias is decreased such that, relative to the baseline scenario, the import share is doubled (HB $\Psi \downarrow$), vacancy posting costs are increased by 25 percent (LME $\xi \downarrow$), and the size of the home country is only 10 percent (CS $\omega \downarrow$).

Table 3. Permanent Effects of Fiscal Devaluation in the Periphery

	Baseline		HB ↓		LME ↓		CS ↓	
Long-Run Changes in	**ΔPIT**	**ΔSSC**	**ΔPIT**	**ΔSSC**	**ΔPIT**	**ΔSSC**	**ΔPIT**	**ΔSSC**
GDP	1.04	0.83	1.02	0.81	1.13	0.92	1.05	0.83
Private Consumption	0.79	0.49	0.61	0.37	0.87	0.52	0.85	0.53
. . . of Optimizers	1.31	0.94	1.02	0.75	1.50	1.08	1.41	1.01
. . . of RoTs	0.27	0.03	0.20	−0.01	0.24	−0.05	0.30	0.05
Private Investment	0.82	0.69	0.64	0.57	0.96	0.82	0.89	0.74
Unemployment Rate	−0.33	−0.39	−0.30	−0.37	−0.35	−0.42	−0.35	−0.40
Per Capita Hours (Private)	0.93	0.54	1.02	0.60	0.99	0.60	0.90	0.52
Total Hours (Private Sector)	1.40	1.09	1.44	1.12	1.48	1.19	1.39	1.08
Average Gross Wages	−1.22	1.16	−1.40	1.04	−1.37	1.03	−1.15	1.21
Average Net Wage Income	4.05	1.71	3.95	1.65	4.18	1.63	4.09	1.74
Unit Labor Costs	−0.96	−1.07	−1.07	−1.15	−1.12	−1.33	−0.92	−1.04
Internat. Competitiveness	1.08	0.71	1.05	0.70	1.13	0.73	1.10	0.73
PIT Rate	−2.98	0.00	−2.98	−0.00	−2.98	0.00	−2.98	−0.00
SSC Rate	0.00	−2.98	0.00	−2.98	0.00	−2.98	0.00	−2.98
Consumption Tax Rate	1.45	0.62	1.49	0.65	1.51	0.68	1.43	0.61
RoE GDP	0.02	0.01	0.03	0.02	0.02	0.01	0.00	0.00

Notes: The table shows deviations of final relative to initial steady-state values in percent (percentage points for rates and ratios). Changes in hours are in percent. International competitiveness is given by foreign prices divided by domestic ones. Columns 1 and 2 show the results for the baseline calibration for PIT and SSC, respectively, while the remaining columns show the results for the robustness analysis of relevant model parameters, which are lower HB = home bias, lower LME = labor market efficiency, and lower CS = country size.

Figure 1. Transition Dynamics (IRFs) for Fiscal Devaluation in the Periphery



The following results stand out. On the one hand, a reduction in labor taxation, regardless of whether on the employees' or employers' side, permanently reduces unit labor costs and induces firms to reduce prices via the marginal costs channel. Lower prices are beneficial to international competitiveness and foster exports. Lower unit labor costs increase labor demand through the creation of additional jobs and/or an increase in the amount of hours worked in the medium run. In the case of the former, unemployment falls. Taken together, aggregate domestic demand for private goods rises. This, plus higher foreign/export demand, increases output, which also fosters the incentive for capital investment. On the other hand, the increase in the consumption tax rate dampens domestic private consumption ceteris paribus, as it makes consumption spending more expensive. The total effect on private consumption turns out to be positive in our simulations, while the size of the impact depends on whether the personal income tax or the social security contribution rate is reduced, because the former increases net labor income by more than the latter.

Focusing on the first-round effects, the impact on consumption differs significantly between the two labor tax instruments. When social security contributions are decreased, the increase in aggregate net wage income is not sufficient to compensate for the increase in consumption costs, and the consumption of rule-of-thumb (RoT) households falls on impact. However, in this case, capital-holding households benefit from higher expected future wealth (due to a permanent rise in output). Under the permanent-income hypothesis, capital-holding households immediately consume more already on impact. When the personal income tax rate is reduced, the increase in net labor income compensates for the higher costs of consumption. As a result, the consumption of RoT households, who spend their entire income each period, already increases on impact (see figure 1). Hence, total private consumption increases more when the personal income tax is reduced.

In the medium to long run, the total impact on consumption and GDP is the result of the mechanisms in the search-and-matching labor market that is incorporated in our model. In what follows, we will restrict ourselves to an intuitive explanation of the different transmission mechanisms after reducing either labor income tax or social security contribution rates. In section 5.1.2 we will shed more light on which assumptions in our model are responsible for these results.

For any given gross wage, a decrease in the labor income tax rate immediately augments net wage income. Workers are then willing to accept lower gross wages in the bargaining process as, everything else equal, they target a certain total net wage income. Thus, private wage claims eventually fall (see the dashed line in figure 1 for the transition and see table 3 for the long run), reducing unit labor costs and increasing the incentive for firms to create jobs and to augment employment. Even though gross wages fall, households' net wage income increases because the gross wage reduction is overcompensated by the reduction in the labor income tax rate. Therefore the consumption of RoT households, who spend their entire per-period income, increases.

A decrease in the social security contribution rate (see the dotted line in figure 1 for the transition and see table 3 for the long run) directly affects the unit labor costs of firms without having to take the detour via the wage-bargaining channel. Again, the fall in unit labor costs increases the incentive for firms to create jobs and augments employment as well as firms' profits. However, it does not have a direct impact on consumers' net income. Since higher employment implies higher chances for unemployed workers to find a job, their fallback position in the wage-bargaining game increases. This augments their reservation wage and, hence, wage claims. Still, unit labor costs fall because wages increase by less than social security contributions fall (unit labor costs are defined as the aggregated gross wage payments including social security payments divided by private-sector output). Relative to a decrease in the personal labor income tax rate, the fall in unit labor costs is stronger on impact and for the first four years.[10] The positive GDP effect is relatively weaker because private consumption demand increases by less.

The different effects of a personal income tax reduction compared with a reduction of social security contributions are crucially driven by the presence of an intensive margin with fully flexible

[10]In the long run, this is reversed because a reduction in the personal labor income tax rate also increases the supply of hours worked (see next paragraph), which gives workers more leeway to accept a lower gross wage while keeping a high enough total net income.

labor hours in our model. This intensive margin allows firms to adjust employment as an input to production differently in the two scenarios. When the personal income tax rate is reduced, the increase in net labor income of workers incites workers to supply more hours. Moreover, it is less costly for firms to produce additional output by increasing working time of those already employed relative to employing more workers, as the latter involves hiring/search costs. On the contrary, when the firms' tax burden is decreased, the increase in net labor income received by workers is lower, and so is the (voluntary) supply of additional working hours. As a result, firms prefer to hire more workers instead of increasing working time. Hence, the total amount of hours worked increases more in case of a personal income tax rate reduction, while unemployment falls more in the case of a reduction in social security contributions (see table 3). In section 5.1.2 we show the importance of incorporating an intensive hours margin in our model.

Interestingly, to achieve budget neutrality, a personal income tax reduction requires a stronger increase in consumption taxes than reducing social security contributions. This is due to the different transmission of the two different measures to wages. As already mentioned, workers accept lower gross wages when reducing the personal labor income tax rate because their net labor income is immediately increased due to the tax rate reduction. On the contrary, gross wage claims are increased after a cut in the social security contribution rate as a result of the bargaining process. Gross wages income is the basis for both personal labor income taxes and social security contributions. Hence, lower wages in the former scenario result in a relatively higher loss of public revenues than a wage increase in the latter scenario, which has to be financed by relatively higher consumption taxation. Summarizing, the model simulations show that from an efficiency perspective—in terms of higher output—lowering labor taxes on the employees' side yields larger gains. This also holds from a redistribution perspective, as it does not harm liquidity-constrained households' consumption. These findings are also in line with Coenen et al. (2012). However, from the perspective of reducing the unemployment rate, a decrease in social security contributions seems more appropriate because, in that case, firms and workers are less inclined to use the intensive margin to increase production. Most of the tax wedge reductions in the past have been a combination of reducing workers' and firms' contribution rates at the same time. The results of such a simulation are presented by the solid line in figure 1. As we see, the impact of this policy mix is a weighted average of both previous simulations. In all cases, spillovers to the core are positive but small, and our results are quite robust to alternative specifications of the periphery's size, trade openess, or labor market efficiency (see table 3).

Even though output and consumption increase more in the case of a personal income tax rate reduction, the effects on welfare are not straightforward because, in this case, the disutility of providing more working hours increases. This is discussed in more detail in section 6.

5.1.2 The Role of Search Frictions and Endogenous Hours Worked

Literature has shown that, in a frictionless world with flexible wages, labor income tax and social security rate reductions are equivalent (see, for example, Lipinska and von Thadden 2012; Engler et al. 2017). This is because, in both cases, wages and labor costs adjust to the same new equilibrium level irrespective of whether the personal income tax or social security tax rates are decreased. In the presence of staggered wage setting, the adjustment of labor costs when lowering the personal labor income tax rate is postponed relative to a reduction in the firms' social security contribution rate, as the latter directly affects labor input costs. Therefore, Engler et al. (2017) find that cutting social security contributions is more beneficial in the short and medium term.[11] In a model with a search-and-matching labor market, the equivalence of personal labor income tax or social security contribution rate reduction in the long run must no longer apply. In these models with a frictional labor market with

[11] Also note that, in this paper, we perform a fully non-linear simulation under perfect foresight that takes into account the fact that the initial and the final steady states are different. This introduces a positive net wealth effect in the long run. Instead, Engler et al. (2017) assume very persistent tax shocks and perform a linear simulation of their model, hence ignoring this wealth effect. Comparing the results shows that, in addition to the different labor market structures, it is also important to actually take into account the final steady state when discussing long-term structural changes.

the standard Nash bargaining over the match surplus—which is also a key feature of our model—personal income taxes and social security contributions affect the bargaining position of workers and firms differently. The following argumentation of the underlying reasons for these results is more formally derived in the appendix.

In general, a reduction in the personal labor income tax rate on the workers' side decreases the workers' outside option in relative terms (i.e., the utility difference between working and not working increases as a result of the policy-induced rise in net wage income). Therefore, workers accept lower wages in the bargaining process, which eventually reduces labor costs for firms. A reduction in the social security contribution rate for firms leaves the workers' outside option nearly unchanged ceteris paribus. However, as the latter increases firms' profits, and the bargaining process determines the share of these profits accruing to the worker, wage claims rise after a reduction in the social security contribution rate. The net effect on labor costs depends on whether the gross wage reduction (after a cut in the personal labor income tax rate) or the direct tax cut (after a reduction in the social security contribution rate) yields stronger effects.

In our model, the outside option of workers depends on unemployment benefits and the expected wage income when finding a job in the public sector. As unemployment benefits relate to net wages, and taxes also have to be paid by public-sector workers, the outside option is affected little after a cut in the personal labor income tax rate. Hence, when ignoring the endogenous evolution of hours worked, we find that cutting social security contributions paid by firms is more beneficial in terms of output, consumption, and employment gains. Table 4 summarizes the results of a simulation in which we keep hours worked fixed at their initial steady-state values. Thus, in this simulation, our model confirms the finding that a cut in the firms' labor tax burden is more beneficial. As table 3 shows, this no longer holds when taking into account the intensive adjustment margin.

The possibility of adjusting the number of hours per worker in the model, next to the decision whether to hire a new worker or not, changes the impact of a reduction in the labor income tax vis-à-vis a reduction in the social security contribution significantly (again, the formal analysis can be retraced in the appendix). As we have

Table 4. Permanent Effects of Fiscal Devaluation in the Periphery with Exogenously Fixed Hours Worked

Long-Run Changes in	ΔPIT	ΔSSC
GDP	0.31	0.40
Private Consumption	0.05	0.06
Private Investment	0.30	0.39
Unemployment Rate	−0.28	−0.37
Per Capita Hours (Private)	0.00	0.00
Average Gross Wages	−0.77	1.43
Average Net Wage Income	3.56	1.43
Unit Labor Costs	−0.72	−0.93
International Competitiveness	0.12	0.15
PIT Rate	−2.98	−0.00
SSC Rate	0.00	−2.98
Consumption Tax Rate	1.81	0.83
RoE GDP	0.01	0.02

Notes: The table shows deviations of final relative to initial steady-state values in percent (percentage points for rates and ratios) after fiscal devaluation for baseline calibration. International competitiveness is given by foreign prices divided by domestic ones.

seen in the model description, workers equate marginal disutility of providing hours with the resulting benefits. A cut in the personal labor income tax rate increases wages/benefits and, thus, the incentives of providing additional hours rise. Furthermore, for firms, it is less costly to extend hours relative to employing an additional worker when extending production (and, thus, labor input), as they can save on search costs. Therefore, augmenting labor input via the hours margin becomes more attractive. This further fosters wage income, consumption, and output. The incentive for workers to provide additional hours increases more after a cut of the personal labor income tax rate than after a cut in the social security contribution rate. The reason is that benefits of providing hours (resulting from the net wage increase) are stronger when the tax burden of workers is decreased, as it has a more direct impact on net wages received by workers (when social security contributions are cut, the wage increase results indirectly through bargaining). Our model simulations show that this effect dominates the job-creation effect such that, when taking into account endogenously determined hours, a cut in the personal labor income tax rate is more beneficial (see table 3).

5.1.3 Alternative Fiscal Instruments to Finance Labor Tax Reductions

We now look at the macroeconomic effects of a decrease in the labor tax wedge when this is financed by alternatively reducing government purchases, C^g, and public employment, N^g. The reduction in the workers' personal income tax rate is taken as the benchmark.

Figure 2 compares a reduction in public purchases and a reduction in public employment to finance the decrease in the tax wedge to a fiscal devaluation in the periphery. Financing a tax wedge reduction via lower government purchases has negative short-run effects on GDP, as it reduces aggregate demand. No adverse effects for consumption of liquidity-constrained households materialize, as they benefit from lower labor taxes. In particular, private consumption of both capital-holding and RoT households increases, as lower labor taxes translate immediately into a higher net wage income while the dampening effect of higher consumption taxes is absent. The improvement in international competitiveness, via lower unit labor costs, increases private consumption and exports. But this is not sufficient to compensate for the loss in public consumption, and GDP declines initially. The negative effect of reduced public consumption is reversed when the labor market improvements resulting from lower labor taxes and the positive wealth effect for optimizers start to materialize. In the medium term, private employment and wages start to increase on the back of higher domestic consumption and exports. This is a result of eventually higher reemployment chances and, therefore, increased fallback utility of (unemployed) workers. Nonetheless, higher private consumption in the long run only slightly compensates for the 1 percentage point loss in public consumption which, in contrast to private consumption, is assumed to entail a full home bias (see table 5). Hence, the private demand-driven GDP increase is dampened, which also implies lower

Figure 2. Transition Dynamics (IRFs) for Alternative Financing Schemes in the Periphery

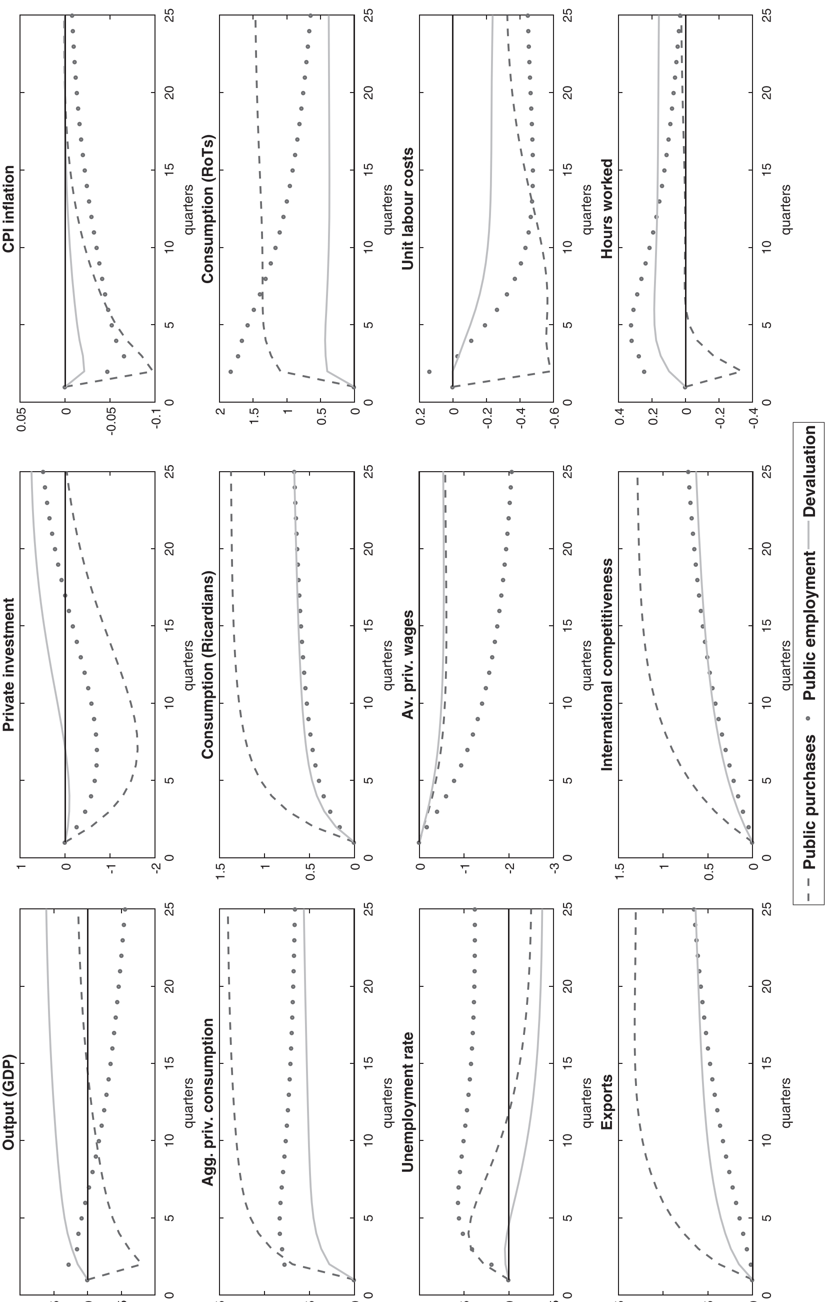


Table 5. Permanent Effects of Financing PIT Reduction with Public Purchases and Employment in Periphery

	Baseline		HB ↓		LME ↓		CS ↓	
Long-Run Changes in	ΔC^g	ΔN^g	ΔC^g	ΔN^g	ΔC^g	ΔN^g	ΔC^g	ΔN^g
GDP	0.57	0.11	0.50	0.02	0.64	0.27	0.59	0.14
Private-Sector Output	0.61	1.61	0.54	1.58	0.69	1.81	0.64	1.62
Private Consumption	1.49	1.29	1.24	1.03	1.63	1.42	1.57	1.38
. . . of Optimizers	1.63	1.98	1.20	1.56	1.87	2.33	1.78	2.14
. . . of RoTs	1.35	0.59	1.29	0.50	1.39	0.51	1.37	0.63
Private Investment	0.16	1.13	−0.16	0.84	0.31	1.39	0.28	1.23
Unemployment Rate	−0.22	0.47	−0.16	0.57	−0.24	0.38	−0.24	0.44
Private Employment	0.22	2.24	0.16	2.27	0.24	2.38	0.24	2.23
Per Capita Hours	0.59	−0.15	0.71	−0.06	0.62	−0.14	0.55	−0.18
Total Hours (Private Sector)	0.89	2.97	0.93	3.10	0.95	3.17	0.88	2.92
Average Gross Wages	−1.26	−2.50	−1.55	−2.85	−1.35	−2.88	−1.16	−2.38
Average Net Wage Income	3.65	1.71	3.48	1.45	3.82	1.57	3.71	1.80
Unit Labor Costs	−0.99	−1.15	−1.17	−1.34	−1.10	−1.51	−0.93	−1.08
Internat. Competitiveness	1.65	1.70	1.64	1.68	1.72	1.79	1.67	1.72
PIT Rate	−2.98	−2.98	−2.98	−2.98	−2.98	−2.98	−2.98	−2.98
Public Purchases	−6.72	0.00	−7.02	0.00	−6.71	0.00	−6.61	0.00
Public Employment Rate	−0.00	−2.72	−0.00	−2.84	0.00	−2.76	−0.00	−2.67
Core GDP	0.03	0.03	0.05	0.05	0.03	0.03	0.00	0.00

Notes: The table shows deviations of final relative to initial steady-state values in percent (percentage points for rates and ratios). Changes in hours are in percent. International competitiveness is given by foreign prices divided by domestic ones. Columns 1 and 2 show the results for the baseline calibration of a PIT reduction financed by a decrease in government consumption and government employment, respectively. The remaining columns show the results for the robustness analysis of relevant model parameters, which are lower HB = home bias, lower LME = labor market efficiency, and lower CS = country size.

effects on overall employment (level and hours worked) in the long run. The GDP-dampening effect of lower public purchases is smaller in countries with a relatively high home bias in private consumption and investment and a relatively low labor market efficiency. Hence, those countries gain more when financing the reduction in the labor tax wedge by lower public purchases.

Two opposing effects are at work, when the labor tax wedge is financed by a decrease in public employment (dotted line in figure 2).

On the one hand, there is a "wage channel" which improves aggregate private output. A decrease in public employment diminishes the probability of finding a job in the public sector, thereby decreasing the workers' fallback utility. Hence, workers will accept lower wages in the private sector beyond what the reduction in the personal income tax rate would entail. This eventually improves unit labor costs and fosters private employment and international competitiveness. The unemployment rate increases even though private employment increases significantly. This is due to the fact that the private sector does not fully absorb the employment reductions in the public sector.

On the other hand, there is a "productivity channel" which lowers aggregate output. As public employees are assumed to positively contribute to private-sector productivity, a reduction in public employment dampens private-sector production capacities. Which of the two effects dominates depends on the relative size of these effects. The wage channel is larger the more workers rely on their fallback utility in the bargaining game. This is the case if the workers' bargaining power vis-à-vis the firms is low. In such a situation, wages fall relatively more after a cut in public employment such that private-sector wage reductions can dominate productivity losses and further improve private-sector output. However, it may have negative distributional consequences in terms of optimizers' and RoTs' consumption behavior (the latter increasing much less) because it shifts income from wages to firms' profits, which belong to optimizers only. In our simulations, the wage channel always dominates and private-sector outcome increases. Hence, we observe an increase in GDP. This is despite the fact that we have defined GDP as the sum of private-sector output and public production evaluated at input costs in line with national accounting (see Stähler and Thomas 2012 for a more detailed discussion).[12]

6. A Welfare Perspective

We are now interested in how to evaluate these reforms in terms of the well-being of the reforming country's population. The advantage of having a theoretical model like ours is that we are able to calculate (household-type-specific) welfare to address this issue. In doing so, we compute the lifetime-consumption-equivalent gain of each type of household in line with Lucas (2003) as a result of the change in fiscal policy. Results are presented in table 6. We first show the welfare difference between the initial and the final steady state and, in a second step, the welfare effects including the transition thereto. The numbers presented in the tables can be interpreted as how much of initial steady-state consumption (in percent) a household would be willing to give up in order to be indifferent between living in the original or in the alternative regime (after the reform). Positive values therefore imply a welfare gain, while negative values signal a welfare loss.

Table 6 shows that a fiscal devaluation via a reduction in the firms' social security contributions always hurts liquidity-constrained consumers. In this case, the gain in net labor income cannot overcompensate for the policy-induced increase in consumption costs. Welfare decreases because, despite lower per capita input of working hours, the loss in consumption utility is too strong to be compensated for. Only because the gains in firms' profits are strong enough to boost optimizers' consumption sufficiently is this measure not detrimental to welfare. Yet, it is still not a Pareto improvement. Once taking into account the transition path, welfare of liquidity-constrained households decreases even more because of the strong

[12]Note furthermore that, in addition to what has just been explained, how the wage and the productivity channels are related is strongly affected by the "efficiency" of the public sector. If the public sector is deemed to be inefficient, this clearly goes in favor of the wage channel. If one believes that the public sector is less efficient in one country than in others, this will strengthen the wage channel in that economy further and, therefore, make this measure relatively more attractive. If private output can be boosted sufficiently, this may overturn the negative GDP effect, which is the case in our simulation.

Table 6. Welfare Gains/Losses in Periphery

Simulation	Baseline			HB ↓			LME ↓			CS ↓		
	ce^o	ce^r	ce^{tot}	ce^o	ce^r	ce^{tot}	ce^o	ce^r	ce^{tot}	ce^o	ce^r	ce^{tot}
In Steady State												
Devaluation (PIT)	1.21	0.18	0.70	0.93	0.11	0.52	1.40	0.14	0.77	1.32	0.21	0.76
Devaluation (SSC)	0.88	−0.03	0.42	0.69	−0.08	0.30	1.01	−0.12	0.45	0.95	−0.01	0.47
PIT and Public Purchases Cuts	1.57	1.29	1.43	1.13	1.22	1.18	1.81	1.33	1.57	1.72	1.31	1.52
PIT and Public Employment Cuts	2.02	0.63	1.32	1.59	0.53	1.06	2.36	0.55	1.45	2.17	0.66	1.42
Including Transition												
Devaluation (PIT)	1.09	0.19	0.64	0.84	0.12	0.48	1.25	0.15	0.70	1.17	0.19	0.68
Devaluation (SSC)	0.81	−0.07	0.37	0.64	−0.11	0.26	0.91	−0.15	0.38	0.86	−0.07	0.39
PIT and Public Purchases Cuts	1.52	1.28	1.40	1.14	1.22	1.18	1.73	1.31	1.52	1.66	1.29	1.47
PIT and Public Employment Cuts	1.13	0.96	1.05	0.71	0.89	0.80	1.39	0.91	1.15	1.27	0.98	1.12

Notes: The table presents steady-state welfare gains/losses after the reform measures simulated in section 5 in terms of how much of initial steady-state consumption (in percent) a household of type $i = o, r$ would be willing to give up in order to be indifferent between living in the original or in the alternative regime (first section, "in steady state"). We then calculate the welfare gains/losses including the transition paths (first section, "including transition"). Total economy-wide welfare gains/losses are defined as $ce^{tot} = (1-\mu)ce^o + \mu^r$, where μ is the share of rule-of-thumb consumers.

initial drop in their consumption on impact. On the contrary, fiscal devaluation by means of a reduction in the workers' personal income tax rate affects welfare positively because, in this case, the negative effect on RoTs' consumption vanishes. Again, when taking into account the transition path, welfare gains are somewhat lower because it takes time to reach the final steady-state values.

Moreover, table 6 also reveals that, in terms of welfare, financing a reduction in the labor tax wedge via a cut in public purchases is superior to a fiscal devaluation, even though the former measure generates significantly lower output gains (see section 5.1.3). This is because reductions in public purchases do not increase consumption costs, while having similar labor market effects. Therefore, the increase in consumption of optimizing and liquidity-constrained consumers is much stronger, translating into higher welfare gains. This also holds when taking into account the transition to the new steady state, again with somewhat lower gains because it takes time to reach the new steady state. A similar argument holds for using a public employment reduction as the financing instrument because, as we have seen in the previous section, this boosts private-sector employment and the increase in aggregate net wage income significantly.

As we have seen above, there are no negative spillovers to the rest of the euro area in terms of output and/or consumption losses. Still, it may be interesting to assess how welfare in the rest of the euro area is affected if one country/region in the euro area reduces its labor tax wedge in a budget-neutral way. Table 7 summarizes the welfare effects in the rest of the euro area—the core in our model. Given the spillovers generated by the tax wedge reduction in the periphery, there are two opposing welfare effects for the rest of the euro area. On the one hand, higher private demand for foreign goods in the periphery increases labor and capital income in the core, ultimately implying higher consumption and, thus, higher welfare there, too. On the other hand, higher output is also produced by augmented labor input, decreasing welfare correspondingly. If the increase in income and, thus, consumption is sufficiently strong to overcompensate for the increase in the disutility of work, households in the core gain. This is the case if private consumption and investment demand in the periphery increases relatively more, which holds more for measures strongly fostering private demand (such as financing the labor tax wedge reduction by public employment

Table 7. Welfare Gains/Losses in Core

Simulation	Baseline			HB ↓			LME ↓			CS ↓		
	ce^o	ce^r	ce^{tot}	ce^o	ce^r	ce^{tot}	ce^o	ce^r	ce^{tot}	ce^o	ce^r	ce^{tot}
In Steady State												
Devaluation (PIT)	−0.04	−0.03	−0.03	−0.03	−0.04	−0.03	0.06	−0.00	0.03	−0.03	−0.05	−0.04
Devaluation (SSC)	0.02	0.01	0.02	0.02	0.01	0.01	0.05	0.02	0.03	−0.01	−0.01	−0.01
PIT and Public Purchases Cuts	0.00	−0.01	−0.01	0.01	−0.02	−0.00	0.18	0.03	0.11	0.05	−0.01	0.02
PIT and Public Employment Cuts	0.07	0.04	0.05	0.09	0.04	0.07	0.14	0.05	0.10	0.13	0.06	0.10
Including Transition												
Devaluation (PIT)	−0.04	−0.03	−0.04	−0.03	−0.03	−0.03	0.06	−0.00	0.03	−0.03	−0.04	−0.04
Devaluation (SSC)	0.02	0.01	0.02	0.02	0.01	0.01	0.05	0.02	0.04	−0.01	−0.01	−0.01
PIT and Public Purchases Cuts	−0.00	−0.01	−0.01	0.02	−0.01	0.00	0.18	0.05	0.12	0.06	0.01	0.04
PIT and Public Employment Cuts	0.05	0.02	0.04	0.06	0.02	0.04	0.10	0.02	0.07	0.07	0.01	0.04

Notes: The table presents steady-state welfare gains/losses after the reform measures simulated in section 5 in terms of how much of initial steady-state consumption (in percent) a household of type $i = o, r$ would be willing to give up in order to be indifferent between living in the original or in the alternative regime (first section, "in steady state"). We then calculate the welfare gains/losses including the transition paths (first section, "including transition"). Total economy-wide welfare gains/losses are defined as $ce^{tot} = (1-\mu)ce^o + \mu^r$, where μ is the share of rule-of-thumb consumers.

cuts). Still, a tax wedge reduction in one country can entail small "beggar-thy-neighbor" effects.

7. Conclusions

Budget-neutral tax wedge reductions are one of the policy priorities in many EMU member states. By means of a New Keynesian DSGE model of a monetary union with a complex labor market structure and a comprehensive fiscal bloc, this paper assessed the macroeconomic and welfare implications of reductions in firms' and workers' labor tax rates financed by different fiscal policy measures. Overall, the paper showed that a reduction in the tax wedge is beneficial in terms of both welfare and output gains. While financing the labor tax wedge reduction by an increase in consumption taxation yielded most favorable output effects, financing it by a reduction in government spending was more welfare enhancing, as the latter does not imply a policy-induced increase in private consumption costs. Extending the insights of existing literature, the paper showed that a reduction in the workers' and not the firms' burden is most beneficial when firms can fully vary the intensive margin of labor demand to adjust to policy changes. When ignoring the intensive hours margin, however, the cost-reducing effect of lower social security contributions for firms dominates, in line with findings in the literature.

Appendix

In this appendix, we provide a more detailed formal analysis of the role of the search-and-matching labor market and of the introduction of an intensive labor hours margin for the impact of workers' personal income tax versus firms' social security contribution reductions on labor supply and demand (cf. section 5.1.2).

Formal Description

First, we will investigate the mechanism driving the labor market effects of a tax wedge reduction in more formal detail. For this purpose, we simplify the model presented in section 3.1 by assuming no liquidity-constrained consumers, $\mu = 0$, and no wage stickiness,

$\theta_w = \theta_w^n = 0$. We will only focus on steady-state comparisons. These assumptions highly simplify the exposition of the argument without loss of generality. Furthermore, we will proceed in four steps for the ease of understanding. Under the simplifying assumptions, equation (6) in steady state becomes

$$\bar{\Omega} = \frac{\xi}{1-\xi} \cdot \frac{1-\bar{\tau}^w}{1+\bar{\tau}^{sc}} \cdot \bar{J}. \tag{15}$$

In a *first step*, let us ignore public employment, the endogenous provision of hours worked, and the time variation in unemployment benefits by exogenously imposing $\bar{p}^g = \kappa^h = 0$, $\bar{l} = 1$, and $\bar{\kappa}^B$ to be fixed at some value. Then, after some algebra, the steady-state wage can be expressed as

$$\begin{aligned}(1-\bar{\tau}^w)\,(1+\bar{\tau}^{sc})\,\bar{w}^p = \xi\,(1-\bar{\tau}^w)\,\left[\bar{x} + \beta(1-s^p)\,\kappa^v \cdot \bar{\theta}^p\right]\\ +\,(1-\xi)\,(1+\bar{\tau}^{sc})\,\bar{\kappa}^B,\end{aligned} \tag{16}$$

where $\bar{\theta}^p = \bar{v}^p/\bar{\tilde{U}}$. Substituting into the job-creation condition, equation (9), and rearranging yields

$$[1-\beta(1-s^p)]\,\frac{\kappa^v}{q(\bar{\theta}^p)} + \beta(1-s^p)\,\xi\,\kappa^v\bar{\theta}^p = (1-\xi)\left[\bar{x} + \frac{1+\bar{\tau}^{sc}}{1-\bar{\tau}^w}\cdot\bar{\kappa}^B\right]. \tag{17}$$

It is straightforward to see that $d\bar{\theta}^p/d\bar{\tau}^w = \frac{1+\bar{\tau}^{sc}}{1-\bar{\tau}^w} \cdot d\bar{\theta}^p/d\bar{\tau}^{sc}$ and, because $\frac{1+\bar{\tau}^{sc}}{1-\bar{\tau}^w} > 1$, it must hold that $|d\bar{\theta}^p/d\bar{\tau}^w| > |d\bar{\theta}^p/d\bar{\tau}^{sc}|$. In words, this implies that, in our simple model, a labor income tax reduction on the workers' side will affect job creation more positively than a reduction of the social security contributions levied on firms. In principle, this already goes in the direction of what we find in our paper. However, from Pissarides (2000, p. 205), we know that, "in general, tax incidence is independent of who pays the tax," while we find that this is not the case here.

In order to solve this alleged contradiction, let us, in a *second step*, allow for time variation in unemployment benefits by assuming that $\kappa_t^B = rrs\,(1-\tau_{t-1}^w)\,w_{t-1}$, as we also do in our model. In this case, we get

$$\begin{aligned}(1-\bar{\tau}^{w})\,(1+\bar{\tau}^{sc})\,(1-(1-\xi)rrs)\,\bar{w}^{p} \\ = \xi\,(1-\bar{\tau}^{w})\,\left[\bar{x}+\beta(1-s^{p})\,\kappa^{v}\cdot\bar{\theta}^{p}\right]\end{aligned} \tag{18}$$

as the steady-state wage, which, after substituting in the job-creation condition and rearranging, yields

$$\begin{aligned}&[1-\beta(1-s^{p})]\,[1-rrs(1-\xi)]\,\frac{\kappa^{v}}{q(\bar{\theta}^{p})}+\beta(1-s^{p})\,\xi\,\kappa^{v}\bar{\theta}^{p} \\ &= (1-\xi)\,(1-rrs)\cdot\bar{x}.\end{aligned} \tag{19}$$

Clearly, it no longer plays a role who pays taxes. As Pissarides (2000, chapter 9) has shown, what matters for job creation is the tax level itself—governed by the parameter rrs in our simplified model—but not who pays the tax. This finding is reconciled in equation (19) and depends on the assumption that unemployment benefits are a fraction of net wages received by workers, which is also the underlying assumption driving the result in Pissarides (2000). The difference to the situation after the first step when assuming fixed unemployment benefits in equation (17) is the following. In the first situation, the relative value of employment over unemployment is affected more by the workers' labor tax rate than by the social security contribution rate due to different effects on the workers' outside option in the bargaining process.[13] Hence, changes in the labor income tax rate will, in this situation, have a larger effect on job creation and it will, then, matter who actually pays the tax. This no longer matters when unemployment benefits are some fraction of net wages.

Even though our model includes time-varying unemployment benefits, simulations still show that results clearly depend on who has to pay the tax. In order to explain why this is the case, let us include public employment into this section's analysis as a *third step*. In this case, we get

$$\begin{aligned}&(1-\bar{\tau}^{w})\,(1+\bar{\tau}^{sc})\left(1-(1-\xi)rrs\left(1-\frac{\beta(1-s^{p})\bar{p}^{g}}{1-\beta(1-s^{g})(1-\bar{p}^{g})}\right)\right)\bar{w}^{p} \\ &= \xi\,(1-\bar{\tau}^{w})\,\left[\bar{x}+\beta(1-s^{p})\,\kappa^{v}\cdot\bar{\theta}^{p}\left(1+\frac{\beta(1-s^{p})\bar{p}^{g}}{1-\beta(1-s^{g})(1-\bar{p}^{g})}\right)\right]\end{aligned}$$

[13]This holds unless $(1+\bar{\tau}^{sc})=(1-\bar{\tau}^{w})^{-1}$, which is not the case in our model and, most likely, not in reality.

$$+ (1-\xi)\frac{\beta(1-s^p)\bar{p}^g}{1-\beta(1-s^g)(1-\bar{p}^g)}(1-\bar{\tau}^w)\,(1+\bar{\tau}^{sc})\,\bar{w}^g, \tag{20}$$

where use has been made of the workers' marginal utility of being employed in the public sector, H_t^g, the latter evaluated at steady state. It is straightforward to see that a higher wage rate in the public sector, $\bar{w}^g$—as well as a higher probability of finding a job in the public sector, $\bar{p}^g$—augments wages workers demand in private-sector wage negotiations. The reason is that the possibility of finding a job in the public sector increases the workers' fallback utility. Substituting the wage resulting from taking into account public employment, equation (20), into the job-creation condition yields

$$\begin{aligned}
&[1-\beta(1-s^p)]\left(1-(1-\xi)rrs\left(1-\frac{\beta(1-s^p)\bar{p}^g}{1-\beta(1-s^g)(1-\bar{p}^g)}\right)\right)\frac{\kappa^v}{q(\bar{\theta}^p)}\\
&\quad+\beta(1-s^p)\,\xi\,\kappa^v\bar{\theta}^p\left(1+\frac{\beta(1-s^p)\bar{p}^g}{1-\beta(1-s^g)(1-\bar{p}^g)}\right) \qquad (21)\\
&=(1-\xi)\,\left(1-rrs\left(1-\frac{\beta(1-s^p)\bar{p}^g}{1-\beta(1-s^g)(1-\bar{p}^g)}\right)\right)\\
&\quad\cdot\bar{x}-(1-\xi)\frac{\beta(1-s^p)\bar{p}^g\cdot(1+\bar{\tau}^{sc})\,\bar{w}^g}{1-\beta(1-s^g)(1-\bar{p}^g)}.
\end{aligned}$$

Formally, we immediately see from equation (21) that $|d\bar{\theta}^p/d\bar{\tau}^w| = 0$, whereas $|d\bar{\theta}^p/d\bar{\tau}^{sc}| > 0$. In words, this means that, when taking into account public employment—still ignoring the endogenous provision of hours worked—a tax wedge reduction using social security contributions levied on firms is more favorable than reducing the workers' labor income tax rate. This is also what we found in the full model when ignoring hours worked.

After the second step in our simplified model, we saw that it makes no difference who pays taxes when assuming time-varying unemployment benefits. The argument is analogous when decreasing the workers' personal income tax rate while, at the same time, taking into account public employment. The reason is that the decrease in $\bar{\tau}^w$ affects steady-state utility of being employed in the private or the public sector in the same direction and by the same relative amount. Hence, the relationship between the workers' steady-state utilities and fallback utilities remains constant. On the contrary, for given public wages and public employment, an increase in social security contributions on the firms' side only reduces job-creation incentives in the private sector, which makes public employment relatively more attractive and increases the workers' fallback utility (in relative terms). Hence, when ignoring the intensive hours margin, the fact that reductions in the firms' social security contributions generate more favorable effects is driven by lowering the relative attractiveness of public-sector employment and, thereby, producing a reduced fallback position of workers.

In a last step, we now also take into account the intensive hours margin. Given that hours in the public sector are assumed to be an exogenous policy variable, we ignore it in the following exposition for the sake of brevity. However, including hours worked in the private-sector bargaining, we need to add $(1-\xi)(1+\bar{\tau}^{sc})\kappa^h/\left(\bar{\lambda}(1+\sigma_h)\right)\bar{l}^{p\,1+\sigma_h}$ to the right-hand-side of equation (20), where $\bar{\lambda}=\left(\bar{c}^{\sigma_c}(1+\bar{\tau}^c)\right)^{-1}$ is households' marginal utility of consumption. Substituting into the job-creation condition, we know that we need to add $-(1-\xi)(1+\bar{\tau}^{sc})/(1+\bar{\tau}^w)\kappa^h/\left(\bar{\lambda}(1+\sigma_h)\right)\bar{l}^{p\,1+\sigma_h}$ there in order to take into account endogenous hours worked in the private sector. Deriving this latter term with respect to $\bar{\tau}^w$ and $\bar{\tau}^{sc}$ implies that, from the perspective of the intensive hours margin, a reduction in the workers' personal income tax rate yields higher incentives for additional job creation as a reduction in firms' social security contributions. The argument is analogous to the one made after the first step, where we had exogenously given unemployment benefits.

Hence, a labor tax wedge reduction by means of a PIT or SSC decrease now entails a tradeoff in the workers' fallback utility. Depending on which instrument is used, it either makes public employment less attractive in relative terms (see equation (21)) or it decreases the relative disutility of labor supply more strongly. Which effect dominates depends on how these two elements in the fallback position of the worker are related. Furthermore, we know from the hours bargaining condition, equation (8) evaluated at steady state, that $|d\bar{l}^p/d\bar{\tau}^w|=\frac{1-\bar{\tau}^{sc}}{1+\bar{\tau}^w}\cdot|d\bar{l}^p/d\bar{\tau}^{sc}|$. This implies that a reduction in the personal income tax rate fosters the provision of (additional) working hours more than a reduction in the firms' social security contribution rate. Therefore, the tradeoff is tilted towards a reduction in the workers' personal income tax rate when taking into account the additional hours margin. Our simulations show that, in the full model, this effect overcompensates for the public employment effect when hours are taken into account.

References

Afonso, A., and P. Gomes. 2014. "Interactions Between Private and Public Sector Wages." *Journal of Macroeconomics* 39 (March): 97–112.

Alesina, A. F., E. L. Glaeser, and B. Sacerdote. 2006. "Work and Leisure in the US and Europe: Why So Different?" In *NBER Macroeconomics Annual 2005*, ed. M. Gertler and K. Rogoff, 1–100 (chapter 1). MIT Press.

Andolfatto, D. 1996. "Business Cycles and Labor-Market Search." *American Economic Review* 86 (1): 112–32.

Attinasi, M.-G., D. Prammer, N. Stähler, M. Tasso, and S. van Parys. 2016. "Budget-Neutral labor Tax Wedge Reductions: A Simulation-Based Analysis for Selected Euro Area Countries." Discussion Paper No. 26/2016, Deutsche Bundesbank.

Balta, N., and J. Delgado. 2009. "Home Bias and Market Integration in the EU." *CESifo Economic Studies* 55 (1): 110–44.

Barro, R. J. 2006. "Rare Disasters and Asset Markets in the Twentieth Century." *Quarterly Journal of Economics* 121 (3): 823–66.

Bassanini, A., and R. Duval. 2006. "The Determinants of Unemployment Across OECD Countries: Reassessing the Role of Policies and Institutions." *OECD Economic Studies* 42 (1): 7–86.

Bilbiie, F. 2008. "Limited Asset Markets Participation, Monetary Policy and (Inverted) Aggregate Demand Logic." *Journal of Economic Theory* 140 (1): 162–96.

Blanchard, O., and J. Galí. 2010. "Labor Markets and Monetary Policy: A New-Keynesian Model with Unemployment." *American Economic Journal: Macroeconomics* 2 (2): 1–30.

Boscá, J. E., A. Díaz, R. Doménech, E. Pérez, J. Ferri, and L. Puch. 2010. "A Rational Expectations Model for Simulation and Policy Evaluation of the Spanish Economy." *SERIEs — Journal of the Spanish Economic Association* 1 (1–2): 135–69.

Boscá, J. E., R. Doménech, and J. Ferri. 2009. "Tax Reforms and Labor-Market Performance: An Evaluation for Spain Using REMS." *Moneda y Credito* 228: 145–88.

———. 2011. "Search, Nash Bargaining and Rule-of-Thumb Consumers." *European Economic Review* 55 (7): 927–42.

———. 2013. "Fiscal Devaluations in EMU." BBVA Working Paper No. 12/11.

Burda, M., and M. Weder. 2002. "Complementarity of Labor Market Institutions, Equilibrium Unemployment and the Propagation of Business Cycles." *German Economic Review* 3 (1): 1–24.

Burgert, M., and W. Roeger. 2014. "Fiscal Devaluation: Efficiency and Equity." Economic Paper No. 542, European Commission (December).

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

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

Christoffel, K., G. Coenen, and A. Warne. 2008. "The New Area-Wide Model of the Euro Area: A Micro-Founded Open-Economy Model for Forecasting and Policy Analysis." Working Paper No. 944, European Central Bank.

Christoffel, K., J. Costain, G. de Walque, K. Kuester, T. Linzert, S. Millard, and O. Pierrard. 2009. "Inflation Dynamics with Labor Market Matching: Assessing Alternative Specifications." Working Paper No. 1053, European Central Bank.

Christoffel, K., K. Kuester, and T. Linzert. 2009. "The Role of Labor Markets for Euro Area Monetary Policy." *European Economic Review* 53 (8): 908–36.

Coenen, G., C. J. Erceg, C. Freedman, D. Furceri, M. Kumhof, R. Lalonde, D. Laxton, J. Lindé, A. Mourougane, D. Muir, S. Mursula, C. de Resende, J. Roberts, W. Roeger, S. Snudden, M. Trabandt, and J. in 't Veld. 2012. "Effects of Fiscal Stimulus in Structural Models." *American Economic Journal: Macroeconomics* 4 (1): 22–68.

Coenen, G., P. McAdam, and R. Straub. 2008. "Tax Reform and Labor-Market Performance in the Euro Area: A Simulation-Based Analysis Using the New Area-Wide Model." *Journal of Economics Dynamics and Control* 32 (8): 2543–83.

Coenen, G., R. Straub, and M. Trabandt. 2013. "Gauging the Effects of Fiscal Stimulus Packages in the Euro Area." *Journal of Economic Dynamics and Control* 37 (2): 367–86.

Colciago, A., T. Ropele, V. A. Muscatelli, and P. Tierlli. 2008. "The Role of Fiscal Policy in a Monetary Union: Are National Automatic Stabilizers Effective?" *Review of International Economics* 16 (3): 591–610.

Cooley, T. F., and E. C. Prescott. 1995. "Economic Growth and Business Cycles." In *Frontiers of Business Cycle Research*, ed. T. F. Cooley, 1–38. Princeton, NJ: Princeton University Press.

CPB. 2013. "Study on the Impacts of Fiscal Devaluation." CPB Netherlands Bureau for Economic Analysis and CAPP in consortium with CASE, CEPII, ETLA, IFO, IFS, HIS. Taxation Paper No. 36, Directorate-General for Taxation and Customs Union, European Commission.

Daveri, F., and G. Tabellini. 2000. "Unemployment, Growth and Taxation in Industrial Countries." *Economic Policy* 15 (1): 47–104.

de Walque, G., O. Pierrard, H. S. Snessens, and R. Wouters. 2009. "Sequential Bargaining in a Neo-Keynesian Model with Frictional Unemployment and Wage Negotiation." *Annales d'Economie et de Statistique* 95–96: 223–250.

Engler, P., G. Ganelli, J. Trevala, and S. Voigts. 2017. "Fiscal Devaluation in a Monetary Union." *IMF Economic Review* 65 (2): 241–72.

European Commission. 2013. "Tax Reforms in EU Member States: Tax Policy Challenges for Economic Growth and Fiscal Sustainability." European Economy Report No. 5/2013.

———. 2014. "Taxation Trends in the European Union." Report.

———. 2016. "The Economic Impact of Selected Structural Reform Measures in Italy, France, Spain and Portugal." European Economy Institutional Paper No. 023 (April).

Farhi, E., G. Gopinath, and O. Itskhoki. 2014. "Fiscal Devaluations." *Review of Economic Studies* 81 (2): 725–60.

Fernández-de-Cordoba, G., J. J. Pérez, and J. L. Torres. 2012. "Public and Private Sector Wages Interactions in a General Equilibrium Model." *Public Choice* 150 (1–2): 309–26.

Forni, L., L. Monteforte, and L. Sessa. 2009. "The General Equilibrium Effects of Fiscal Policy: Estimates for the Euro Area." *Journal of Public Economics* 93 (3–4): 559–85.

Gadatsch, N., K. Hauzenberger, and N. Stähler. 2016. "Fiscal Policy During the Crisis: A Look on Germany and the Euro Area with GEAR." *Economic Modelling* 52 (Part B): 997–1016.

Gadatsch, N., N. Stähler, and B. Weigert. 2016. "German Labor Market and Fiscal Reforms 1999-2008: Can They Be Blamed for Intra-Euro Area Imbalances?" *Journal of Macroeconomics* 50 (December): 307–24.

Galí, J., J. D. Lopez-Salido, and J. Vallés. 2007. "Understanding the Effects of Government Spending on Consumption." *Journal of the European Economics Association* 5 (1): 227–70.

Gertler, M., L. Sala, and A. Trigari. 2008. "An Estimated Monetary DSGE Model with Unemployment and Staggered Nominal Wage Bargaining." *Journal of Money, Credit and Banking* 40 (8): 1713–64.

Gomes, S., P. Jacquinot, and M. Pisani. 2016. "Fiscal Devaluation in the Euro Area: A Model-Based Analysis." *Economic Modelling* 52 (Part A): 58–70.

Kaufmann, C. 2016. "Optimal Fiscal Substitutes for the Exchange Rate in a Monetary Union." Discussion Paper No. 44/2016, Deutsche Bundesbank.

Keane, M. P. 2011. "Labor Supply and Taxes: A Survey." *Journal of Economic Literature* 49 (4): 961–1075.

Krebs, T. 2003. "Growth and Welfare Effects of Business Cycles in Economies with Idiosyncratic Human Capital Risk." *Review of Economic Dynamics* 6 (4): 846–68.

Langot, F., L. Patureau, and T. Sopraseuth. 2014. "Fiscal Devaluation and Structural Gaps." Working Paper No. 508, Banque de France.

Le Blanc, J., A. Porpiglia, F. Teppa, J. Zhu, and M. Ziegelmeyer. 2014. "Household Saving Behavior and Credit Constraints in the Euro Area." Discussion Paper No. 16/2014, Deutsche Bundesbank.

Leeper, E. M., T. B. Walker, and S. C. S. Yang. 2010. "Government Investment and Fiscal Stimulus." *Journal of Monetary Economics* 57 (8): 1000–1012.

Lipinska, A., and L. von Thadden. 2009. "Monetary and Fiscal Policy Aspects of Indirect Tax Changes in a Monetary Union." Working Paper No. 1097, European Central Bank.

———. 2012. "On the (In)Effectiveness of Fiscal Devaluations in a Monetary Union." Finance and Economics Discussion Series No. 2012-71, Board of Governors of the Federal Reserve System.

Lucas, R. 2003. "Macroeconomic Priorities." *American Economic Review* 93 (1): 1–14.

Meghir, C., and D. Phillips. 2010. "Labor Supply and Taxes." In *Dimensions of Tax Design: The Mirrlees Review*, ed. S. Adams, 202–74. Oxford: Oxford University Press.

Merz, M. 1995. "Search in the Labor Market and the Real Business Cycle." *Journal of Monetary Economics* 36 (2): 269–300.

Moyen, S., N. Stähler, and F. Winkler. 2019. "Optimal Unemployment Insurance and International Risk Sharing." *European Economic Review* 115 (June): 144–71.

Nickell, S., L. Nunziata, and W. Ochel. 2005. "Unemployment in the OECD Since the 1960s. What Do We Know?" *Economic Journal* 115 (500): 1–27.

Obstfeld, M. 1994. "Evaluating Risky Consumption Paths: The Role of Intertemporal Substitutability." *European Economic Review* 38 (7): 1471–86.

Ohanian, L., A. Raffo, and R. Rogerson. 2008. "Long-Term Changes in Labor Supply and Taxes: Evidence from OECD Countries, 1956–2004." *Journal of Monetary Economics* 55 (8): 1353–62.

Otrok, C. 2001. "On Measuring the Welfare Cost of Business Cycles." *Journal of Monetary Economics* 47 (1): 61–92.

Pappa, E. 2009. "The Effects of Fiscal Shocks on Employment and the Real Wage." *International Economic Review* 50 (1): 217–44.

Petrongolo, B., and C. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39 (2): 390–431.

Picos-Sánchez, F., and A. Thomas. 2015. "A Revenue-Neutral Shift from SSC to VAT: Analysis of the Distributional Impact of 12 EU-OECD Countries." *FinanzArchiv: Public Finance Analysis* 71 (2): 278–98.

Pissarides, C. 2000. *Equilibrium Unemployment Theory*. 2nd Edition. Cambridge, MA: MIT Press.

———. 2009. "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?" *Econometrica* 77 (5): 1339–69.

Prescott, E. C. 2004. "Why Do Americans Work So Much More than Europeans?" Working Paper No. 10316, National Bureau of Economic Research.

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

Smets, F., and R. Wouters. 2003. "An Estimated Stochastic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association* 1 (5): 1123–75.

———. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97 (3): 586–606.

Stähler, N., and C. Thomas. 2012. "FiMod — A DSGE Model for Fiscal Policy Simulations." *Economic Modelling* 29 (2): 239–61.

Thimann, C. 2015. "The Microeconomic Dimension of the Eurozone Crisis and Why European Politics Cannot Solve Them." *Journal of Economic Perspectives* 29 (3): 141–64.

Using Payment System Data to Forecast Economic Activity*

Valentina Aprigliano, Guerino Ardizzi, and Libero Monteforte
Banca d'Italia

Payment systems trace economic transactions; they could therefore be considered important indicators of economic activity. This paper describes the monthly data available on Italy's retail settlement system and selects some of them for nowcasting and short-term forecasting. Using a mixed-frequency factor model based on a large-scale data set to predict Italian GDP and its main components, the contribution of payment system flows to improving forecasting accuracy is found to be non-negligible. Moreover, the timeliness of the data improves nowcasting accuracy throughout the quarter.

JEL Codes: C53, E17, E27, E32, E37, E42.

1. Introduction

Ever since the global recession, interest in new macroeconomic forecasting tools, especially those based on monetary and financial information, has been increasing. On the back of the developments of computational tools for storing and elaborating large-scale data sets, analysts are focusing on the pursuit of new, timely, and reliable information in order to improve the forecasting ability in real time.

Data on payment instruments (checks, credit transfers, direct debits, payment cards) could represent a unique source of information for short-term forecasting of economic activity, as they trace economic transactions. This link was already clear at the beginning of the last century, when Irving Fisher described the seminal equations of the quantitative theory of money, writing: "Such

*This paper represents the authors' personal opinions and does not reflect the views of the Bank of Italy. We are grateful to Luca Arciero, Alessandro Barbarino, Claudio Impenna, Massimiliano Marcellino, Antonio Rua, Roberta Zizza, Francesco Zollino, and two anonymous referees for their useful comments. We also thank the participants who attended the CIRET Conference, Copenhagen 2016, RCEA Workshop, Rimini 2016, IAAE Conference, Sapporo 2017.

elementary equations mean that the money paid in any transaction is the equivalent of the goods bought at the price of sale" (Fisher 1912). Moreover, the importance of payments, together with banking and asset markets, for understanding how monetary economics works has recently received renewed attention from the growing body of research called New Monetarist Economics (see Williamson and Wright 2010 and Schneider and Piazzesi 2015). Nevertheless, the use of payment data for macroeconomic forecasting has only recently been exploited.

Studies for Canada (Galbraith and Tkacz 2007, 2009, 2018), Portugal (Esteves 2009; Duarte, Rodrigues, and Rua 2017), Denmark (Carlsen and Storgaard 2010), and the United States (Barnett et al. 2016) find that payment transactions can help with nowcasting and with forecasting GDP and private consumption in the short term.

It is worth noting that this strand of the literature concentrates on payment cards (as in Esteves 2009; Carlsen and Storgaard 2010; Duarte, Rodrigues, and Rua 2017; Galbraith and Tkacz 2018; and Barnett et al. 2016, who also consider checks) and mainly targets GDP and consumption, thus offering only a partial view of the whole system of payments and of economic activity.

To the best of our knowledge, this paper is the first attempt to assess the ability of retail payment data to make accurate short-term forecasts, focusing on Italy and drawing on a broad set of aggregated payments. We target both GDP and its main domestic components (households' consumption, HC, gross fixed investments, GFI, and value added in the service sector, VAS) and we use a comprehensive set of payment instruments, including credit transfers, checks, direct debits, and debit cards. This approach allows us to conduct a more robust empirical application than similar studies for other countries, which mostly consider a subset of payment instruments and macroeconomic aggregates. Our data are recorded electronically through clearing and settlement circuits managed by the Bank of Italy (for retail transactions, BI-Comp; for wholesale payments including customer transactions, BI-REL up to May 2008 and TARGET2-Bank of Italy subsequently), and they are not revised because they are recorded without errors by construction.[1] Another noteworthy

[1]Our data are extracted from the payment system infrastructures, while in some structural analysis data are collected by means of surveys or diaries (see Bagnall et al. 2016).

feature of these data is their timeliness. Indeed, the payment data are available on a daily basis with a short delay.

We show that there is a close correlation between retail payment series (hereinafter PS) and the main macroeconomic aggregates. In the empirical application, we select the indicators by means of LASSO (least absolute shrinkage and selection operator) and then we set up a mixed-frequency dynamic factor model to predict the quarter-on-quarter growth of GDP and its main components by using a large-scale monthly data set, which includes standard business cycle indicators other than payment data; we perform out-of-sample forecasting simulations, including and excluding PS. The contribution of PS turns out to be appreciable and promising throughout the forecasting horizon considered (from one quarter backwards up to two quarters ahead). Moreover, the timeliness of the payment data lets us refine the nowcasting throughout the quarter in real time.

The rest of the paper is organized as follows. Section 2 gives an overview of the payment system in Italy. Section 3 introduces the data, providing some descriptive evidence on the relationship between PS and the main macroeconomic aggregates. Section 4 deals with the empirical application: subsection 4.1 shows that LASSO picks PS amongst the first fourteen predictors to be included in the model (out of the fifty indicators compiled earlier); subsection 4.2 describes the forecasting exercise and the results. Section 5 concludes.

2. Overview of the Payment System for the Italian Economy

A payment system is the set of instruments, rules, procedures, and technologies used to settle money transfers among economic agents.

We can distinguish between *wholesale payments* and *retail payments*. The former typically involve the banking system handling large-value payments (interbank transactions), usually connected with financial markets flows and refinancing operations with national central banks; the latter refer to transactions within the circuit of individuals and firms and closely related to economic activity (Padoa-Schioppa 2004).

Before the launch of the euro in 1999, the payment system in Europe was highly fragmented. Monetary union posed the problem of harmonizing the infrastructure to transfer money among economic operators to foster financial and commercial integration. The broad-based reorganization of the payment system proved mostly effective for wholesale payments, operated by two area-wide systems: TARGET2 (Trans-European Automated Real-Time Gross Settlement Express Transfer System; hereinafter T2), provided by the Eurosystem, and EURO1, privately owned. The Bank of Italy, along with Deutsche Bundesbank and Banque de France, helped to develop T2, which settles the majority of wholesale transfers on a gross-real-time basis.

As for retail payments,[2] the system is not yet fully integrated. However, since 2014 the Single Euro Payments Area (hereinafter SEPA) has strongly promoted the standardization and interoperability of different national clearing and settlement retail systems. The Bank of Italy manages the BI-Comp clearing system, which works in accordance with the rules of SEPA. BI-Comp clears domestic payments on a multilateral net basis. These payments can be settled both in BI-Comp and in T2. In fact, due to urgency and for security reasons, banks may prefer to settle customers' payments in T2. This retail branch of T2 is named T2-retail from here on in.

Retail non-cash payments settled through BI-Comp and T2-retail add up to €5 trillion on a yearly basis (about 60 percent of the total value of retail payments in Italy—80 percent if we only consider electronic payments, excluding postal pre-printed processed and other paper-based credit transfers), about three times higher than the nominal value of GDP.[3] Although in Italy cash payments

[2]The retail payments system generally uses the clearing and settlement mechanism (CSM), in which one or more operators perform clearing (i.e., transmission, matching, confirmation of payments, and calculation of a settlement position) and settlement (completion of the payment).

[3]Unlike cash payments, which amount to immediate transfers of value between the payer and the payee through bank notes and coins, non-cash payments are exchanges of funds through accounts. It follows that the relationship between the payer and the payee is mediated by authorized institutions (such as banks and postal offices), which actually process the payment before settling the transaction. This brings us to a crucial distinction depending on the party submitting the payment order: that between *credit-based* instruments (i.e., credit transfers, card payments), submitted by the payer, and *debit-based* instruments (i.e., direct

remain the most frequently used instrument for retail payments (about 80 percent of the *number* of retail payments), new information and communication technologies have encouraged non-cash payments, mainly those processed through electronic devices, allowing greater flexibility and customization. If we consider the *value* of the transactions, on average per year,[4] the share of cash payments shrinks to about 45 percent of consumer-to-business transactions (for instance, point-of-sales, or POS, purchases) and to less than 10 percent of all transactions, including business-to-business payments. It is worth noting that the payment data recorded in BI-Comp include ATM cash withdrawals, which may represent a good proxy of the transactions paid with cash (see Schmiedel, Kostova, and Ruttenberg 2012). Data on cash withdrawals also allow us to take into account the long-run trends on consumer payment habits (cash versus non cash) within our forecasting approach. For the Italian economy, the percentage of cash withdrawals on the total value of payments is fairly stable (below 5 percent) in the time span considered in the empirical application. In the same period, the figure is also broadly stable for the euro area. Figure 1 shows the monthly gross flow of retail payments settled through BI-Comp and T2-retail in Italy.

The sharp decrease observed in 2014 for BI-Comp stems from changes in the customer payment landscape following the migration to SEPA.[5] Some participants reconsidered the routing policies for their customer payments, and they ultimately opted for SEPA-compliant automated clearinghouses other than BI-Comp and T2. However, T2-retail has been less affected, because it meets some specific customers' demands concerning urgency and assurance of payments.

debits, checks), submitted by the payee. Credit transfers are the most common credit-based instrument, while debit-based instruments include direct debits, card payments, and checks. For a detailed description of payment instruments, see http://www.bancaditalia.it/compiti/sispaga-mercati/strumenti-pagamento/index.html?com.dotmarketing.htmlpage.language=1 and *The Payment System: Payments, Securities and Derivatives, and the Role of the Eurosystem*, by the European Central Bank.

[4]These data refer to 2014.

[5]SEPA represents a harmonization of procedures and platforms in the euro zone for processing retail credit transfers and direct debits.

Figure 1. Total Payments Settled in BI-Comp and TARGET2-Retail (values)

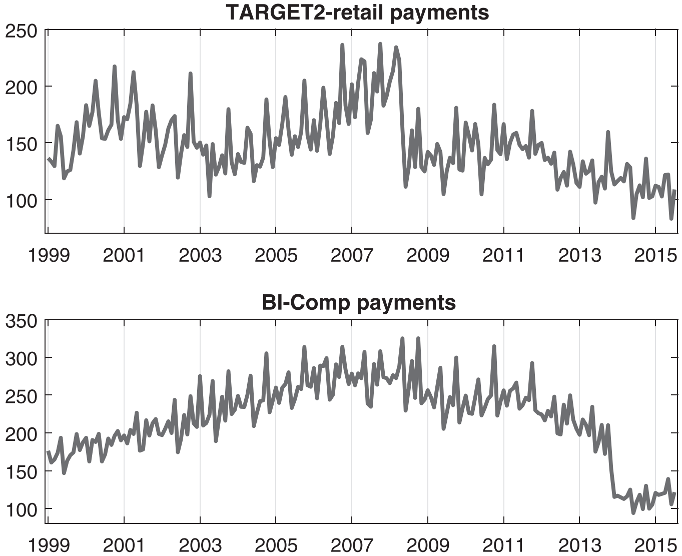
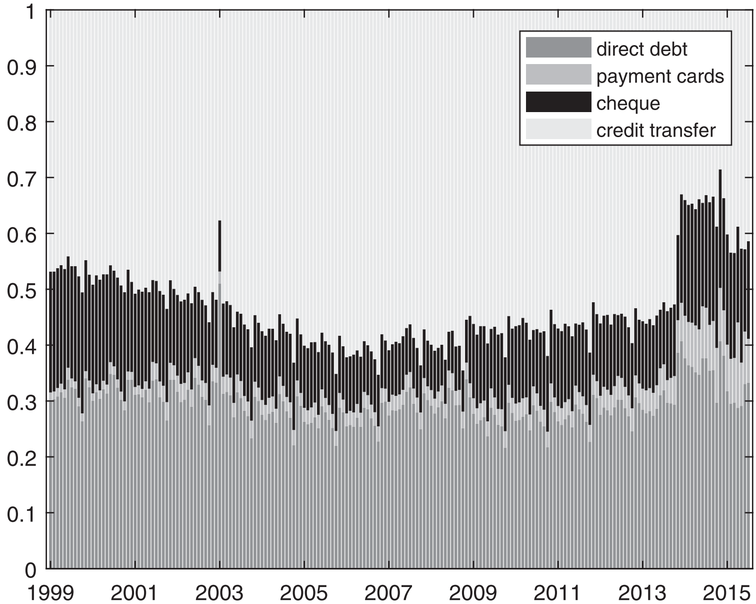


Figure 2 depicts the share of transactions enabled by different payment instruments over the total gross flows settled in BI-Comp. Credit transfers represent the largest share (almost 55 percent);[6] that of direct debits is also sizable (29 percent) since they provide a very neat solution to managing recurrent payments (e.g., utility bills, mortgage payments). Payment cards are far less used in Italy, accounting for only 3 percent of the total non-cash payments; cards are used for low-value transactions (€50–€100 for POS; €100–€200 for ATM).

However, a large share of credit card payments are recorded as direct debits, since credit card statements are often charged to a payer's current account.

[6]For business purposes credit transfers are the most popular and suitable payment instrument, accounting for 80 percent of the total value of business-to-business transactions; direct debits and checks account for 10 percent, respectively.

Figure 2. Total Payments Settled in BI-Comp

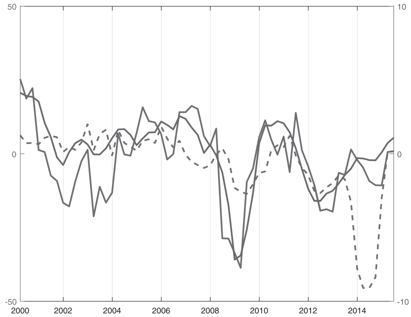


3. Payments Data and Macroeconomic Aggregates

In this section we look at some descriptive statistics which prove the close relationship between the main macroeconomic aggregates (such as GDP, value added in the service sector, private consumption, and gross fixed investments) and the payment flows. In our analysis, we focus on retail payment instruments since they are used to settle commercial transactions as consumer-to-business and business-to-business payments. We find clear empirical evidence underpinning the role of payment time series in tracking economic activity.

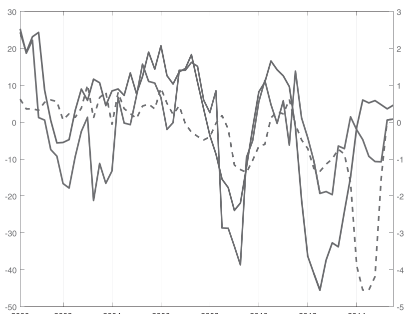
We rely on the BI-Comp and T2-retail systems, described in section 2, to collect timely and high-frequency data. Data are collected by the central bank for the individual transactions of the payment systems. These systems close every day, therefore the observation error is basically nil. We do not use individual payments (which eventually could be relevant for a big data analysis) but instead use the aggregated monthly time series stored in Bank of

Figure 3. BI-Comp (dashed line) and T2-Retail (solid gray line; left-side axis) Flows Compared with Macroeconomic Aggregates (solid black line; right-side axis): Year-on-Year Percentage Changes

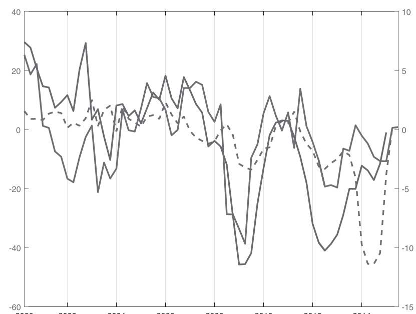


(a) Payment flows and GDP

(b) Payment flows and households' consumption



(c) Payment flows and value added in service sector



(d) Payment flows and gross fixed investment

Italy's database, which is publicly available.[7] The monthly observations are simply the sum of the nominal value of all the individual payments recorded in the month.

The annual growth of GDP and of its main components co-moves closely with the annual growth of payment flows settled through BI-Comp and T2-retail (see figure 3).[8] The big drop at the end of the

[7]Data may be found at https://www.bancaditalia.it/statistiche/tematiche/statistiche-sistema-pagamenti/index.html.

[8]BI-Comp and T2-retail nominal flows are divided by the GDP national account deflator in order to obtain a measure of the volume of transactions. In

Table 1. Correlation between Payment Flows and Macroeconomic Indicators (year-on-year percentage changes)

Payment Flows[a]	GDP	Private Consumption	Gross Fixed Investment	Value Added Service Sector
BI-Comp	71.8	70.5	74.3	68.4
T2-Retail	79.4	67.7	64.0	74.3
BI-Comp + T2-Retail	90.7	82.5	82.3	85.9
Other Indicators				
M2	−38.1	−11.5	−20.4	−45.6
Industrial Production	92.8	68.7	74.3	79.8
Business Confidence[b]	68.5	74.9	67.3	64.3
Consumer Confidence[c]	9.9	40.0	14.3	19.7

[a] Contemporaneous unconditional correlations are computed on the sample 2000:Q1–2012:Q4 in order to exclude the break caused by the new standard SEPA.
[b] Economic Sentiment Indicator (ESI) provided by Istat.
[c] Consumer confidence survey by Istat.

sample proves that BI-Comp was more affected than T2-retail by the launch of SEPA in October 2014.[9] Some interesting results also emerge from the correlation matrix (see table 1).[10] The correlation between PS and the target variables is valuable and similar to the one shown by other indicators such as industrial production and business confidence, usually adopted in short-term forecasting. The

the empirical application (section 4), we consider nominal value, but we include the harmonized consumer price index as a variable to control for the price effect instead of deflating the payment series.

[9] The launch of SEPA has entailed the switching of some domestic credit transfers and direct debits from BI-Comp into STEP2, which is the pan-European infrastructure managed by a private body (Interbank Society for Automation). Data on payment flows in STEP2 are not publicly available. Nonetheless, removing the outlier corresponding to the launch of SEPA turns out to be suited to preserving the predictive ability of BI-Comp.

[10] We consider the contemporaneous correlation.

results in table 1 are consistent with the picture of the payment instruments composition in Italy. Credit transfers and direct debits are the payment tools most relevant in terms of values; they are also more closely correlated with the target variables than payment cards and checks.

4. Empirical Application

4.1 Selection of the Targeted Predictors

We compiled N = 50 variables, which provide a fairly complete picture of economic activity, including indicators of industrial and service activity (industrial production, electricity consumption, freight truck, business confidence), households' consumption (retail sales of goods and services, consumers' confidence), financial indexes, and credit flows to firms other than time series from the payment systems T2-retail and BI-Comp. All the variables are seasonally adjusted by monthly seasonal dummies; as for PS, we also remove some outliers pinned down as the peaks and troughs larger than 1.5 times the standard deviation of the time series. These outliers are replaced by the mean of the neighboring observations. The payment series refer to the nominal value of the transactions.[11] We also use nominal estimates of GDP and of its main components. The construction of the appropriate deflator for both PS and the macroeconomic aggregates would require a specific analysis, therefore we decided to include the price index[12] in the final model in order to control for the price effect.

Factor models provide a parsimonious way of handling such a large-scale data set. Nevertheless, some recent studies have demonstrated that using a large number of variables to estimate the common factors may worsen the forecasting performance (see Boivin and Ng 2006). Therefore, we use LASSO (for a review of the LASSO estimator, see Tibshirani 1996 and Hastie, Tibshirani, and Friedman 2009) to screen the variables depending on their ability to anticipate our targets, in the same spirit as Bai and Ng (2008), who use both

[11] The number of transactions is also available, but in this paper we focus on the values.

[12] We use the harmonized index of consumer prices for Italy.

hard and soft thresholding methods to select the so-called targeted predictors. Moreover, the selection made by LASSO provides initial evidence of whether PS can be considered good predictors of the short-term evolution of economic activity, compared with other workhorse indicators for conjunctural analysis.

The monthly indicators are transformed into quarterly variables and then the LASSO regression is performed. LASSO selects the regressors by solving the minimization problem

$$min\left(\sum_{T}^{t=1}(y_t - \sum_{j=1}^{N}\beta_j X_{jt})^2\right) \quad s.t. \sum_{j=1}^{N}|\beta_j| \leq \tau, \tag{1}$$

where τ is the tuning parameter. LASSO selects n_L targeted variables, among all N collected:

$$\hat{L}_{n_L} = \{j \in \{1, 2, \ldots, N\} : |\hat{\beta}_{Lj}| > 0\}. \tag{2}$$

We revise the LASSO selection by introducing and discarding a few variables as documented below.[13] With respect to the GDP model, we get rid of the credit flows and PS-total (the sum of BI-Comp and T2-retail); we deem the latter redundant once BI-Comp and T2-retail are included. As for households' consumption model, LASSO would pick BI-Comp and PS-total, but we find it reasonable to replace the latter (which carries redundant information if BI-Comp is already included) with T2-retail; we discard consumers' survey on "future personal economic situations" and two series from the business surveys ("expected level of orders" and "expected level of liquidity" in the consumer goods sector) which track industrial activity more effectively. In the model for gross fixed investments, the "expected level of production" in the intermediate goods sector is less volatile than the "expected level of orders." We include BI-Comp and T2-retail otherwise discarded by LASSO in both the model for investments and for value added in the service sector. In the latter we also include the price index, excluded by LASSO. Table 2 shows the variables finally included in our information set, $\hat{I}_n$.

[13] As a robustness check, we estimated three alternative models. More precisely, in the first model LASSO picks the targeted predictors and PS are in the pool of candidates; in the second model PS are not among the candidates; and the third model expands the second model with PS.

Table 2. The Final Model ($\hat{I}_n$)

Indicators	GDP	HC	GIF	VAS
Total Electricity Consumption	No	Yes	Yes	Yes
Industrial Production	Yes	Yes	Yes	Yes
Business Climate	No	No	Yes	Yes
CCS—Future General Economic Situations	No	No	Yes	Yes
CCS—Future Personal Economic Sit.	No	No	Yes	Yes
CCS—Unemployment Exp.	No	No	Yes	Yes
CCS—Saving Opportunities, Next Twelve Months	Yes	Yes	Yes	Yes
CCS—Households Balance Sheet	Yes	Yes	Yes	Yes
CCS—Current Saving Opportunities	No	No	Yes	No
BCS—Current Level of Orders (Intermediate Goods)	No	No	Yes	No
BCS—Current Level of Production (Int. Goods)	No	No	Yes	No
BCS—Expected Level of Production (Int. Goods)	No	No	Yes	Yes
BCS—Exp. Level of Orders (Int. Goods)	No	No	No	No
BCS—Future General Economic Sit. (Int. Goods)	No	No	No	No
BCS—Exp. Level of Liquidity (Int. Goods)	No	No	No	No
BCS—Current Level of Liquidity (Int. Goods)	No	No	No	No
BCS—Current Level of Orders (Investment Goods)	No	No	No	No
BCS—Current Level of Production (Inv. Goods)	No	No	No	No
BCS—Current Level of Liquidity (Inv. Goods)	No	No	No	No
BCS—Exp. Level or Orders (Inv. Goods)	No	No	No	No
BCS—Exp. Level of Production (Inv. Goods)	No	No	No	No
BCS—Future General Economic Sit. (Inv. Goods)	No	No	No	No
BCS—Exp. Level of Liquidity (Inv. Goods)	No	No	No	No
BCS—Current Level of Orders (Consumer Goods)	No	No	No	No
BCS—Current Level of Production (Cons. Goods)	No	No	No	No
BCS—Exp. Level of Orders (Cons. Goods)	Yes	No	Yes	No
BCS—Exp. Level of Production (Cons. Goods)	No	No	No	No
BCS—Future General Economic Sit. (Cons. Goods)	Yes	No	No	No
BCS—Exp. Level of Liquidity (Cons. Goods)	No	No	No	No
BCS—Current Level of Liquidity (Cons. Goods)	No	Yes	No	No
Current Accounts Deposits (Stock)	No	No	No	No
Credit Flows to Firms	No	No	No	No
HICP	Yes	Yes	Yes	Yes
FTSE Italy (Banks)	No	No	No	No
FTSE Italy (Insurance)	No	No	No	No
FTSE Italy (Transport)	No	No	No	No
PMI Services—Business Activity	No	No	No	No
PMI Services—New Business	Yes	No	No	No
PMI Manufacturing	No	No	No	No
PMI Manufacturing—Output	No	No	No	No
PMI Manufacturing—New Orders	Yes	Yes	Yes	Yes
PMI Manufacturing—Employment	No	No	No	No
PMI Manufacturing—New Export Orders	No	Yes	No	No
Freight Truck	Yes	No	No	Yes
Retail Trade—Goods	Yes	Yes	No	Yes
Retail Trade—Services	No	Yes	No	No
BI-COMP	Yes	Yes	Yes	Yes
TARGET RETAIL	Yes	Yes	Yes	Yes
Payments System—Total	No	No	No	No

We implement a large-scale dynamic factor model on $\hat{I}_n$, which is more suitable when the cross-section size of the data set, n, is valuable with respect to the number of observations, T. Working with large data sets proves a robust strategy for detecting whether PS contribute to improving accuracy in the estimation of economic growth, since they are compared to many other indicators, often used to track the short-term dynamics of the economy.

4.2 Forecasting Exercise

4.2.1 The Model

Let $\mathbf{X}_t$ be the n-vector of observable monthly variables selected earlier and $\mathbf{X}$ the $T \times n$ matrix. They are driven by q common factors, $\mathbf{f}_t = [f_{1t}, \ldots, f_{qt}]^{'}$, and by n idiosyncratic components, $\xi_t = [\xi_{1t}, \ldots, \xi_{nt}]^{'}$, which are assumed to be uncorrelated. The $n \times q$ matrices Λ_i, for $i = 1 \ldots s$ lags, are the common-factor loadings. The dynamic factor model

$$\begin{aligned}\mathbf{X}_t &= \Lambda_s(L)\mathbf{f}_t + \xi_t \\ &= \Lambda_0\mathbf{f}_t + \Lambda_1\mathbf{f}_{t-1} + \Lambda_2\mathbf{f}_{t-2} + \cdots + \Lambda_s\mathbf{f}_{t-s} + \xi_t\end{aligned} \tag{3}$$

can be mapped to the static model

$$\mathbf{X}_t = \mathbf{D}\mathbf{F}_t + \xi_t, \tag{4}$$

where $\mathbf{F}_t = [\mathbf{f}_t^{'}, \mathbf{f}_{t-1}^{'}, \ldots \mathbf{f}_{t-s}^{'}]'$ has the dimension $r = q(s+1)$ and can be represented as a VAR(l) process,[14]

$$\mathbf{F}_t = \mu + \Psi_1\mathbf{F}_{t-1} + \cdots + \Psi_l\mathbf{F}_{t-l} + \mathbf{u}_t. \tag{5}$$

As stated in Bai and Ng (2007), $\mathbf{F}_t$ is driven by $q < r$ common shocks if $\mathbf{u}_t = \mathbf{A}\epsilon_t$, where ϵ_t is a q-vector of mutually orthogonal shocks with variance equal to one and the $r \times q$-matrix $\mathbf{A}$ has rank q; then $E(\mathbf{u}_t\mathbf{u}_t^{'}) = \Sigma_u = \mathbf{A}\Sigma_\epsilon\mathbf{A}^{'}$ has reduced rank q. We estimate $\mathbf{A}$ as in

[14] We set the order l of the VAR equal to 4.

Marcellino and Schumacher (2010): given the OLS estimate $\hat{\Sigma}_u$ and its eigenvalue decomposition $\mathbf{MPM}^{'}$, let $\mathbf{M}_*$ be the $r \times q$-matrix of the first q eigenvectors and $\mathbf{P}_*$ the diagonal matrix of the corresponding eigenvalues; then $\hat{\mathbf{A}} = \mathbf{M}_*\mathbf{P}_*^{-1/2}$ and the reduced-rank estimate of Σ_u is equal to $\Sigma_u^r = \hat{\mathbf{A}}\hat{\mathbf{A}}^{'}$. We follow Bai and Ng (2002, 2007) to set the number of static ($r = 4$) and dynamic ($q = 1$) factors and then we estimate factors' space $\mathcal{G}(F_t) = \overline{\text{span}}(F_{1t}, \ldots, F_{rt})$ by principal components extracted from the balanced monthly data set. More specifically, we extract from the covariance matrix of $\mathbf{X}$ the first r eigenvalues and the corresponding $n \times 1$ eigenvectors, $\mathbf{V}_i$ for $i = 1, \ldots, r$, and then we compute the $T \times r$ matrix of the common static factors $\mathbf{F} = \mathbf{XV}$. By construction, the covariance matrix of the common static factors is the diagonal matrix of the first r eigenvalues, while $\hat{\Sigma}_\xi$ is a diagonal matrix whose entries are extracted by the diagonal of the covariance matrix of $\mathbf{X}-\mathbf{F}\hat{\mathbf{D}}^{'}$, where $\hat{D}$ is the OLS estimate of the $n \times r$ matrix of static-factors loadings.

The quarterly growth rate of the target variable, y_{t_q} with t_q labeled by a multiple of the last month of each quarter (i.e., $t_q = 3, 6, \ldots, 3\lceil T/3\rceil$), is projected by an unrestricted mixed-data sampling (MIDAS) model on the monthly information $\mathcal{G}(F)$ (see Ghysels, Santa-Clara, and Valkanov 2004 for an extensive treatment of the MIDAS model and refer to Foroni, Marcellino, and Schumacher 2015 for the unrestricted MIDAS model). For each quarter t_q we have $m = 3$ values of the monthly regressors; therefore, we apply the $L^{j/m}$ operator to obtain regressors lagged by j months with respect to the quarter. Put formally,

$$\begin{aligned} y_{t_q} &= c + \beta_0\mathbf{F}_{t_q} + \beta_1\mathbf{F}_{t_q-1/m} + \cdots + \beta_p\mathbf{F}_{t_q-p/m} + \epsilon_{t_q} \\ &= c + \beta(L^{1/m})\mathbf{F}_{t_q} + \epsilon_{t_q}, \end{aligned} \tag{6}$$

where the loadings β_j for $j = 1, \ldots, p$ have dimension $1 \times r$ and are estimated by a simple OLS. We choose $p = 3$. By way of example, let us suppose Q1 is the current quarter and $t_q = 3\lceil t/3\rceil$. Therefore, the estimate of y_{t_q} will rely on information on December, January, February, and March. Equations (4)–(6) are cast in a state-space form. To include the quarterly growth rate of the target variable into the state-space framework, we construct a monthly series y_t, where $y_t \equiv y_{t_q}$ when $t \equiv t_q$, and is missing otherwise. If $l \geq p$, the state equation is

$$\mathbf{X}_t = \underbrace{\begin{bmatrix}\mathbf{D} & 0_{n\times r(l-1)+1}\end{bmatrix}}_{\mathbf{DD}} \cdot \begin{bmatrix}\mathbf{F}_t^* \\ \\ y_t\end{bmatrix} + \mathbf{e}_t, \tag{7}$$

where

$$\mathbf{F}_t^* = \begin{cases}[\mathbf{F}_t', \mathbf{F}_{t-1}', \ldots \mathbf{F}_{t-l+1}']' & \text{if } l \geq p \\ [\mathbf{F}_t', \mathbf{F}_{t-1}', \ldots \mathbf{F}_{t-p+1}']' & \text{otherwise}\end{cases} \tag{8}$$

and $\mathbf{e}_t = [\xi_t, 0]' \sim N(0_{(n+1)\times 1}, \mathbf{R})$ with $\mathbf{R} = \begin{bmatrix}\Sigma_\xi & 0 \\ 0 & 0\end{bmatrix}$; the transition equation is

$$\underbrace{\begin{bmatrix}\mathbf{I}_{rl} & & 0_{rl\times 1} \\ & & \\ -\beta_0 & 0_{1\times r(l-1)} & 1\end{bmatrix}}_{H} \cdot \begin{bmatrix}\mathbf{F}_t^* \\ \\ y_t\end{bmatrix} = \underbrace{\begin{bmatrix}\begin{bmatrix}\mathbf{Psi} \\ \mathbf{I}^*\end{bmatrix} & 0_{rl\times 1} \\ & \\ \begin{bmatrix}\beta_1 \cdots \beta_p\end{bmatrix} & 0_{1\times r(l-p)+1}\end{bmatrix}}_{G} \cdot \begin{bmatrix}\mathbf{F}_{t-1}^* \\ \\ y_{t-1}\end{bmatrix}$$

$$+ \underbrace{\begin{bmatrix}\mu \\ 0_{r(l-1)\times 1} \\ c\end{bmatrix}}_{d} + \begin{bmatrix}\mathbf{u}_t \\ 0_{r(l-1)\times 1} \\ \epsilon_t\end{bmatrix}, \tag{9}$$

where

$$\begin{bmatrix}\mathbf{Psi} \\ \mathbf{I}^*\end{bmatrix} = \begin{bmatrix}\Psi_1 & \Psi_2 & \cdots & \Psi_l \\ \mathbf{I}_r & & 0_{r\times r(l-1)} & \\ \vdots & & & \\ 0_{r\times r(l-2)} & & \mathbf{I}_r & 0_{r\times r}\end{bmatrix},$$

while if $l < p$, the state equation is

$$\mathbf{X}_t = \begin{bmatrix}\mathbf{D} & 0_{n\times r(p-1)+1}\end{bmatrix} \cdot \begin{bmatrix}\mathbf{F}_t^* \\ \\ y_t\end{bmatrix} + \mathbf{e}_t \tag{10}$$

and the transition equation is

$$\begin{bmatrix} \mathbf{I}_{rp} & 0_{rp\times 1} & \\ -\beta_0 & 0_{1\times r(p-1)} & 1 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{F}_t^* \\ y_t \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \mathbf{Psi} \\ \mathbf{I}^* \end{bmatrix} & 0_{rl\times r(p-l)+1} \\ \multicolumn{2}{c}{0_{r(p-l)\times rp+1}} \\ [\beta_1 \cdots \beta_p] & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{F}_{t-1}^* \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \mu \\ 0_{r(p-1)\times 1} \\ c \end{bmatrix} + \begin{bmatrix} \mathbf{u}_t \\ 0_{r(p-1)\times 1} \\ \epsilon_t \end{bmatrix} \tag{11}$$

with $\begin{bmatrix} \mathbf{u}_t \\ \underline{0} \\ \epsilon_t \end{bmatrix} \sim N(\underline{0}, \mathbf{Q})$ and $\mathbf{Q} = \begin{bmatrix} \Sigma_u^r & \underline{0} \\ \underline{0} & \sigma_\epsilon^2 \end{bmatrix}$. We implement the Kalman recursions (Kim and Nelson 1999) to extract the smoothed state variable $\mathbf{FF}_t = [\mathbf{F}_t^{*'} \quad y_t]'$. Let T_ν be the vintage of the monthly series and h the forecasting horizon in terms of quarters; then the smoothed state variable is the estimate at time t, for $t = 1, 2, \ldots, T_\nu + 3h$, based on information up to $T = T_\nu + 3h + 1$, i.e., $\mathbf{FF}_{t|T}$. This is very appreciable because we use the latest information to infer the dynamics of the state variable. In fact, when $t = T_\nu+1, \ldots, T_\nu+3h$, all observations are missing and they are given no weight by the Kalman filter, which basically forecasts the factors as elaborated in Giannone, Reichlin, and Small (2008). In practice, we impose infinite variance to the i-th idiosyncratic component if the i-th observable variable is missing. In so doing, the weight of the missing observation will fade when computing the Kalman gain, which projects the **X**s onto the space spanned by the idiosyncratic factors. In particular, we are interested in forecasting the target variable at time $T_{q_\nu} + h$, where T_{q_ν} is the quarter corresponding to the month T_ν:

$$y_{T_{q_\nu}+h} = c + \beta(L^{1/m})\mathbf{F}_{T_{q_\nu}+h|T} + \epsilon_{T_{q_\nu}+h}. \tag{12}$$

4.2.2 Out-of-Sample Simulation

We conduct an out-of-sample forecasting exercise at different horizons ($h = -1, 0, 1, 2$)[15] to assess how the forecasting performance of our models changes depending on whether we include or exclude PS from the set of regressors.

We run a pseudo real-time simulation; therefore, we use the latest available vintage of data and we cut it period by period, being careful to replicate the missing values' pattern at the end of the sample.[16] This data set ranges from January 2000 to November 2015, and it is balanced on August 2015.

The exercise is carried out on two different time intervals. The first estimation sample goes from January 2000 to April 2008 and expands until the last balanced date (i.e., August 2015). The current period is assumed to be two months after the balanced date, therefore the first *pseudo* vintage is June 2008 and we nowcast 2008:Q2 and make two-steps-ahead forecasts for 2008:Q3 and 2008:Q4. We end up with the first sample of forecasting errors, which is 2008:Q2–2015:Q2. The second estimation sample goes from January 2000 to July 2011 and the sample of the forecasting errors is 2011:Q3–2015:Q2.

The benchmark model includes all the variables listed in the column "Final Model" of table 2 and is compared with the model replacing PS with the first variable discarded by LASSO for each target $\hat{I}_{-PS}$.[17] In fact, we consider two benchmark models: the first includes only T2-retail such as payment series ($\hat{I}_{T2}$) while the second

[15]The forecasting horizon depends on which month of the quarter is taken as the *pseudo* vintage. For instance, in the first month of quarter t_q, we backcast ($h = -1$) the target variable of the previous quarter not yet released by the national statistical office (Istat); we also nowcast ($h = 0$) and forecast one quarter and two quarters ahead ($h = 1, 2$).

[16]It is reasonable that the final estimate of GDP is related more closely to the payment series as well as to the final revised versions of other macroeconomic variables. If the latter were in real time, then the payment series may have been given an unfair advantage. Therefore, by making everything observed as their final revised values, we prevent this risk.

[17]Electricity consumption for GDP; business climate for HC; orders' expectations—intermediate goods for GFI; households' current saving convenience for VAS. As for GFI and VAS, PS are initially discarded by LASSO and we reintroduce them in place of orders' expecations—intermediate goods and households' current saving convenience, respectively.

Table 3. Relative RMSFE of the Model including only T2-Retail

		Backcast	Nowcast	Forecast One Step	Forecast Two Steps
HC	2008:Q2–2015:Q2	1.10	1.14	1.15	1.14
	2011:Q3–2015:Q2	0.97	0.98	1.04	1.06
GFI	2008:Q2–2015:Q2	1.00	0.93	1.06	1.07
	2011:Q3–2015:Q2	0.80	0.82	1.17	1.12
VAS	2008:Q2–2015:Q2	1.00	1.08	1.11	1.12
	2011:Q3–2015:Q2	1.01	1.03	0.99	0.98
GDP	2008:Q2–2015:Q2	0.74	0.86	1.04	1.11
	2011:Q3–2015:Q2	1.28	1.10	0.91*	0.94

Notes: This table shows the relative RMSFE of the model including only T2-retail ($\hat{I}_{T2}$) vis-à-vis the model replacing payment series with the first variable discarded by LASSO for each target ($\hat{I}_{-PS}$). For GDP, the Diebold-Mariano test (with HAC variance estimators) is computed: the null hypothesis is that the out-of-sample errors of the competing models are equal. * shows p-values < 0.05, ** p-values < 0.01, and *** p-values < 0.001.

includes both T2-retail and BI-Comp ($\hat{I}_{T2,BC}$). Tables 3 and 4 show the root mean square forecasting error (RMSFE) of the competing models $\hat{I}_{-PS}$ relative to the RMSFE of the two benchmark models. A figure greater than one means that PS improve the forecasting performance. PS improve the forecasting accuracy broadly.[18] In general, information on retail payments plays a role when the targets are projected one quarter and two quarters ahead, irrespective of the time interval, and it tracks well the dynamics of households' consumption.[19] In particular, we record a forecasting gain throughout the horizons in $\hat{I}_{T2,BC}$.

[18]We compared the forecast accuracy of non-nested models (when we exclude the payment variable we include the first excluded indicator in LASSO); therefore, we used the Diebold-Mariano test (with heteroskedasticity and autocorrelation consistent, HAC, variance estimators) of forecast accuracy. We computed this test for GDP forecasts, but we did not find a clear pattern of statistical significance. We interpret this finding as a result of the small out-of-sample window in our forecasting application.

[19]Regarding the three alternative specifications described in footnote 13, the results on the role of payment series in forecasting are mixed. However, these models generally worsen the forecasting performance compared with the benchmark $\hat{I}_{T2,BC}$ (the results are available on request).

Table 4. Relative RMSFE of the Model including both T2-Retail and BI-Comp

		Backcast	Nowcast	Forecast One Step	Forecast Two Steps
HC	2008:Q2–2015:Q2	1.23	1.21	1.13	1.10
	2011:Q3–2015:Q2	1.09	1.07	1.09	1.06
GFI	2008:Q2–2015:Q2	1.10	1.07	1.24	1.22
	2011:Q3–2015:Q2	0.80	0.82	1.17	1.12
VAS	2008:Q2–2015:Q2	1.08	1.13	1.09	1.09
	2011:Q3–2015:Q2	1.10	1.08	1.01	1.01
GDP	2008:Q2–2015:Q2	0.93	0.98	1.07	1.09
	2011:Q3–2015:Q2	1.79	1.58**	1.05	1.01

Notes: This table shows the relative RMSFE of the model including both T2-retail and BI-Comp ($\hat{I}_{T2,BC}$) vis-à-vis the model replacing payment series with the first variable discarded by LASSO for each target ($\hat{I}_{-PS}$). For GDP, the Diebold-Mariano test (with HAC variance estimators) is computed: the null hypothesis is that the out-of-sample errors of the competing models are equal. * shows p-values < 0.05, ** p-values < 0.01, and *** p-values < 0.001.

Model $\hat{I}_{T2,BC}$ performs better than $\hat{I}_{T2}$; however, the latter also yields some valuable results. T2 shrinks the RMSFE for backcasting and nowcasting GDP by 28 percent and 10 percent, respectively, during the turmoil sparked by the sovereign debt crisis; in the same period, we gain 17 percent of predictive accuracy when forecasting investments one step ahead.

Some remarkable results come out of $\hat{I}_{T2,BC}$. The forecasting performance improves for HC and GFI throughout the horizons when we consider the longest sample. As for activity in the service sector, in the same period PS lowers the RMSFE for nowcasting by 13 percent. The most notable outcome regards the backcast and the nowcast of GDP during the last four years. In this case, using PS lowers the RMSFE by 79 percent and 58 percent, respectively.

These results bear out the suitability of PS to making predictions of real activity. The model we used, which belongs to the class of models with large data sets, strengthens our claim further. Indeed, within a factor model framework the marginal contribution of the single indicator to the covariance of the common components typically fades as the cross-section dimension of the data set becomes sizable.

Table 5. Relative RMSFE for the Second and the Third Month of the Quarter vis-à-vis the First Month (percentage values)

	2005:Q2–2015:Q2	
	$\hat{I}_{T2,BC}$	$\hat{I}_{-PS}$
First Month		
Second Month	−5.9	−2.9
Third Month	−21.0	−18.0

Note: The figure is the ratio between the RMSFE in both the second and the third month of the quarter and in the first month.

One of the most appealing characteristics of the payment system data is their timeliness, since they are available in the reference month. Therefore, we can assess how much we benefit from the monthly real-time information on PS to improve the nowcast during the quarter. This experiment is conducted for GDP only. PS are assumed to be the most timely regressors of the GDP model, while the other explicative variables are two months late with respect to the reference period. Using the same *pseudo* real-time simulation exercise, we update the nowcast each month of the quarter and we observe how the RMSFE changes. We compare the nowcast made in the first month of the quarter with the nowcast made in the second and in the third month by $\hat{I}_{T2,BC}$ and $\hat{I}_{-PS}$, where the latter replaces PS with the total electricity consumption, which is as timely as PS. In so doing, we lose two-thirds of the observations, so we use a longer sample, from 2005:Q2 to 2015:Q2. In table 5 we observe a monotonic improvement of the forecasting accuracy from the first to the last month of the quarter for both models, as expected. However, $\hat{I}_{T2,BC}$ performs somewhat better than $\hat{I}_{-PS}$. When we are in the second month of the quarter, the nowcast by $\hat{I}_{T2,BC}$ is 5.9 percent more accurate than the nowcast made in the first month (2.9 percent by $\hat{I}_{-PS}$). The accuracy improves further in the third month of the quarter (the nowcast is 21 percent more precise than in the first month—18 percent by $\hat{I}_{-PS}$).

4.2.3 Observation Weights

One of the most common remarks made about factor models concerns the possibility of disentangling the contribution of the observable variables to the forecasts. The variance of y_{t_q} is explained by the common factors, as shown in (6), which are not given any definite economic meaning. As showed in Koopman and Harvey (2003), the output of the Kalman recursions can be used to measure the weights attached to the observable variables in $\mathbf{X}_j$ when forecasting y_{t_q}, for $j = 1, \ldots, T_\nu + 3h + 1$. The smoothed state vector can be expressed as the weighted sum:

$$\mathbf{FF}_{t|T} = \sum_{j=1}^{T} w_j(\mathbf{FF}_{t|T})\mathbf{X}_j, \tag{13}$$

where the weights $w_j(\cdot)$ are a function of the state vector. Each month t, the weights are computed by the backward recursions (for $j = t-1, t-2, \ldots, 1$) and the forward recursions (for $j = t, t+1, \ldots, T$) introduced in Koopman and Harvey (2003).[20] Let us define $\mathbf{L}_t = \mathbf{Z} - \mathbf{K}_t\mathbf{DD}$, $N_t = \mathbf{DD}^{'}\mathbf{S}_t^{-1}\mathbf{DD} + \mathbf{L}_t^{'}\mathbf{N}_t\mathbf{L}_t$ for $t = T, \ldots, 1$ and $\mathbf{N}_T = 0$; $\mathbf{C}_t = \mathbf{DD}^{'}\mathbf{J}_t - \mathbf{Z}^{'}\mathbf{N}_t\mathbf{K}_t$ where $\mathbf{J}_t = \mathbf{S}_t^{-1} + \mathbf{K}_t^{'}\mathbf{N}_t\mathbf{K}_t$ for $t = T, \ldots, 1$. The weights for filtering are given by the following recursions:

$$w_j(\mathbf{FF}_{t|t-1}) = B_{t,j}\mathbf{K}_j, \quad B_{t,j-1} = B_{t,j}\mathbf{Z} - w_j(\mathbf{FF}_{t|t-1})\mathbf{DD} \tag{14}$$

for $j = t-1, \ldots, 1$ with $B_{t,t-1} = \mathbf{I} - \mathbf{P}_{t|t-1}\mathbf{N}_{t-1}$, while the weights for smoothing are given by

$$w_j(\mathbf{FF}_{t|T}) = (I - \mathbf{P}_{t|t-1}\mathbf{N}_{t-1})w_j(\mathbf{FF}_{t|t-1}), \quad j < t \tag{15}$$

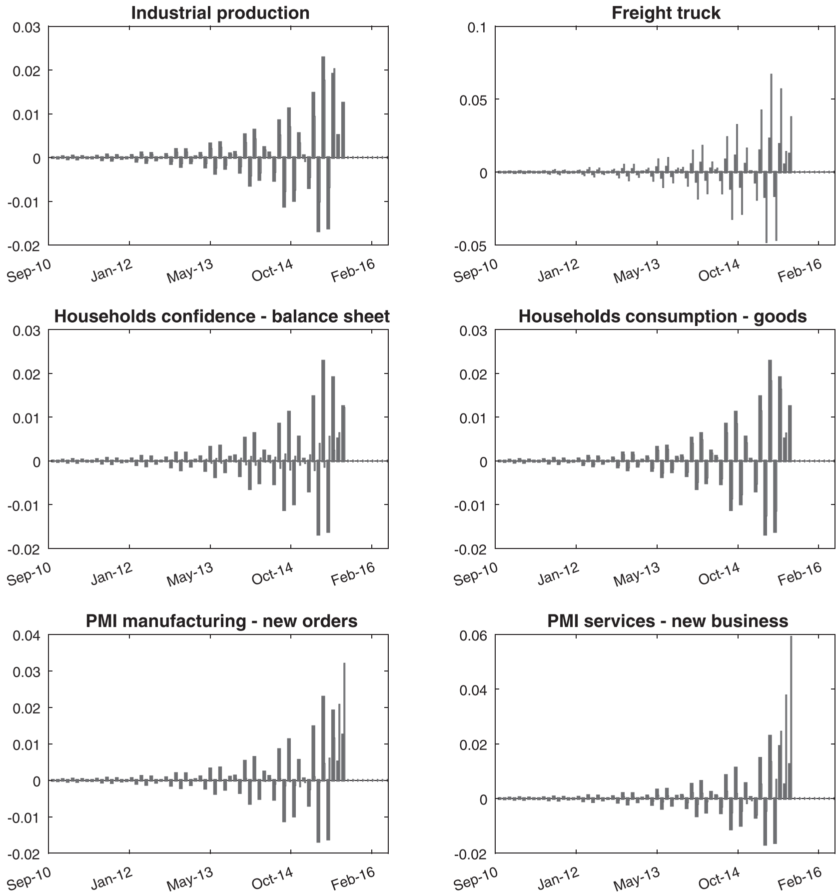
$$w_j(\mathbf{FF}_{t|T}) = B_{t,j}^*\mathbf{C}_j \quad B_{t,j+1}^* = B_{t,j}^*\mathbf{L}_j^{'}, \quad j = t, \ldots, T \tag{16}$$

with $B_{t|t}^* = P_{t|t-1}$.

Figure 4 shows the weights attached to T2-retail (black bars) and to some other indicators (gray bars), which track the short-term evolution of economic activity very well (such as the

[20]A detailed description of the algorithms may be found in Koopman and Harvey (2003, p. 1322).

Figure 4. Weights for Nowcasting (PS: black bar; other variable: gray bar)



industrial production index; freight truck; households' confidence about their own balance sheet; households' consumption of goods; manufacturing PMI—new orders; and service PMI—new business). Let us assume that the current month is October 2015. Therefore, we need to anticipate GDP growth 2015:Q4 (nowcasting).

The weights of PS are comparable to those of industrial production and of households' consumption of goods; the latest available data on PS help make up for the missing information on industrial production and households' consumption at the end of the sample. PS have sizable weights compared with those of the qualitative surveys (PMIs and households' confidence about the balance sheet); this means that past information on PS is more effective than that on very timely and cyclical indicators in anticipating GDP growth.

5. Conclusions

Our findings show that payment data track economic activity. We look at different aggregates of payment system flows in Italy, together with other indicators usually adopted in macroeconomic forecasting, and we see that they maintain some additional information content. We start from a large database of short-term monthly indicators, and LASSO selects the payments jointly with other standard business cycle indicators (e.g., industrial production and business surveys), for both GDP and households' consumption. Moreover, an out-of-sample forecasting application using a mixed-frequency factor model shows that the model including retail payment flows generally outperforms the one based on standard short-term indicators only, in terms of forecasting accuracy. The results are shown not only for GDP but also for consumption, investments, and value added in the service sector. In order to disentangle the contribution of the observable variables to the forecasts of GDP, we estimate the weights proposed in Koopman and Harvey (2003), and we find that the weights attached to PS are comparable to those of some of the most important short-term indicators generally used to track economic activity. Using the mixed-frequency feature of our model, we show that the timeliness of the payment data improves the forecast accuracy during the quarter, more than other comparable short-term indicators.

The way we pay is changing following the digitalization of retail payments, and this process fosters the production of big data, which the analysts can rely on. Payment data indicators used in this model are the aggregate of a huge number of transactions. However, the forecasting ability of these indicators paves the way for future research exploring the big data structure of the individual transactions.

References

Bagnall, J., D. Bounie, K. P. Huynh, A. Kosse, T. Schmidt, S. Schuh, and H. Stix. 2016. "Consumer Cash Usage: A Cross-Country Comparison with Payment Diary Survey Data." *International Journal of Central Banking* 12 (4): 1–61.

Bai, J., and S. Ng. 2002. "Determining the Number of Factors in Approximate Factor Models." *Econometrica* 70 (1): 191–221.

———. 2007. "Determining the Number of Primitive Shocks in Factor Models." *Journal of Business and Economic Statistics* 25 (1): 52–60.

———. 2008. "Forecasting Economic Time Series Using Targeted Predictors." *Journal of Econometrics* 146 (2): 304–17.

Barnett, W., M. Chauvet, D. Leiva-Leon, and L. Su. 2016. "Nowcasting Nominal GDP with the Credit-Card Augmented Divisia Monetary Aggregates." Working Papers Series in Theoretical and Applied Economics No. 201605, University of Kansas, Department of Economics.

Boivin, J., and S. Ng. 2006. "Are More Data Always Better for Factor Analysis?" *Journal of Econometrics* 132 (1): 169–94.

Carlsen, M., and P. E. Storgaard. 2010. "Dankort Payments as a Timely Indicator of Retail Sales in Denmark." Technical Report.

Duarte, C., P. M. Rodrigues, and A. Rua. 2017. "A Mixed Frequency Approach to the Forecasting of Private Consumption with ATM/POS Data." *International Journal of Forecasting* 33 (1): 61–75.

Esteves, P. S. 2009. "Are ATM/POS Data Relevant When Nowcasting Private Consumption?" Technical Report.

Fisher, I. 1912. *The Purchasing Power of Money.* New York: Macmillan Company.

Foroni, C., M. Marcellino, and C. Schumacher. 2015. "Unrestricted Mixed Data Sampling (MIDAS): MIDAS Regressions with Unrestricted Lag Polynomials." *Journal of the Royal Statistical Society: Series A* (*Statistics in Society*) 178 (1): 57–82.

Galbraith, J. W., and G. Tkacz. 2007. "Electronic Transactions as High-Frequency Indicators of Economic Activity." Technical Report.

———. 2009. "A Note on Monitoring Daily Economic Activity via Electronic Transaction Data." CIRANO Working Paper No. 2009s-23.

———. 2018. "Nowcasting with Payments System Data." *International Journal of Forecasting* 34 (2): 366–76.

Ghysels, E., P. Santa-Clara, and R. Valkanov. 2004. "The MIDAS Touch: Mixed Data Sampling Regression Models." CIRANO Working Paper No. 2004s-20.

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.

Hastie, T. J., R. J. Tibshirani, and J. H. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer Series in Statistics. New York: Springer.

Kim, C.-J., and C. R. Nelson. 1999. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications,* Vol. 1. The MIT Press.

Koopman, S. J., and A. Harvey. 2003. "Computing Observation Weights for Signal Extraction and Filtering." *Journal of Economic Dynamics and Control* 27 (7): 1317–33.

Marcellino, M., and C. Schumacher. 2010. "Factor MIDAS for Nowcasting and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP." *Oxford Bulletin of Economics and Statistics* 72 (4): 518–50.

Padoa-Schioppa, T. 2004. "Shaping the Payment System: A Central Bank's Role." Speech delivered at the Bank of Korea's Conference on Payment Systems, Seoul, May 13. Available at https://www.ecb.europa.eu/press/key/date/2004/html/sp040513_1.en.html.

Schmiedel, H., G. Kostova, and W. Ruttenberg. 2012. "The Social and Private Costs of Retail Payment Instruments: A European Perspective." ECB Occasional Paper No. 137.

Schneider, M., and M. Piazzesi. 2015. "Payments, Credit and Asset Prices." 2015 Meeting Paper No. 133, Society for Economic Dynamics.

Tibshirani, R. 1996. "Regression Shrinkage and Selection Via the Lasso." *Journal of the Royal Statistical Society: Series B (Methodological)* 58 (1): 267–88.

Williamson, S., and R. Wright. 2010. "New Monetarist Economics: Models." In *Handbook of Monetary Economics*, Vol. 3, ed. B. M. Friedman and M. Woodford, 25–96 (chapter 2). Elsevier.

Monetary Policy during Financial Crises: Is the Transmission Mechanism Impaired?*

Nils Jannsen,[a] Galina Potjagailo,[a,b] and Maik H. Wolters[a,c,d]
[a]Kiel Institute for the World Economy
[b]University of Kiel
[c]University of Jena
[d]IMFS at Goethe-University Frankfurt

The effects of monetary policy during financial crises differ substantially from those in normal times. Using a panel VAR for twenty advanced economies, we show that monetary policy has larger and quicker effects during financial crises on output and inflation, and also on various other macroeconomic variables like credit, asset prices, uncertainty, and consumer confidence. The effects on output and inflation are particularly strong during the acute phase of financial crises when the economy is also in recession, while they are weaker during the subsequent recovery phase. We find differences in the size and the timing of monetary policy actions during the global financial crisis of 2008/09 across countries that may have contributed to the different macroeconomic performance across countries.

JEL Codes: C33, E52, E58, G01.

1. Introduction

During the global financial crisis that started in 2007, many central banks eased monetary policy aggressively in order to alleviate financial market distress, boost output, and stabilize inflation. Monetary

*Financial support by Deutsche Bundesbank Regional Office in Hamburg, Mecklenburg-West Pomerania and Schleswig-Holstein is greatly acknowledged. We thank Martin Ademmer, Jens Boysen-Hogrefe, Kai Cartensen, Georgios Georgiadis, Peter Griep, Philip Lowe, John Murray, Volker Wieland, and participants at various seminars and conferences for helpful comments and discussions. Author e-mails: nils.jannsen@ifw-kiel.de; gpotjagailo@gmail.com; maik.wolters@uni-jena.de.

policy was largely successful in mitigating financial market distress, but output growth and inflation remained lower than expected in many advanced economies and the recoveries were widely perceived as disappointingly sluggish (see, e.g., Pain et al. 2014). These observations led to a broad-based debate on whether the transmission channels of monetary policy were impaired due to the global financial crisis and whether monetary policy is in general less effective during financial crises and their aftermath (see, e.g., Bouis et al. 2013).

The relevance of this debate goes beyond the central banks' general need for assessing the effectiveness of their policies. It is crucial for defining a policy mix that can stabilize the economy during financial crises. If monetary policy is less effective in such crises, the questions arise whether larger monetary stimulus needs to be provided in order to achieve the desired effects; whether other policies, such as fiscal policy, need to be used more extensively; or whether sluggish recoveries after financial crises must be tolerated given that there is little scope for monetary policy. In this regard, the question of monetary policy effectiveness during financial crises is also relevant for the assessment of undesirable side effects, such as excessive risk-taking and asset price bubbles, that can occur if monetary policy remains highly expansionary for an extended period of time (see, e.g., Rajan 2005; Altunbasa, Gambacorta, and Marques-Ibanez 2014; Jiménez et al. 2014).

Financial crises exhibit several characteristics that can influence the transmission of monetary policy: high financial market distress, macroeconomic volatility and uncertainty, low confidence of market participants, and substantial balance sheet adjustments of firms and households. All these adverse characteristics might impair the transmission of monetary policy. In particular, banks might be unwilling to extend their credit supply, because they face higher credit default risk and because they need to adjust their balance sheets after previous losses (Bouis et al. 2013; Buch, Buchholz, and Tonzer 2014; Valencia 2017). Also, financial crises are typically preceded by asset price bubbles and credit and consumption booms and, as a consequence, are typically followed by deleveraging of households and firms and increasing risk aversion (Reinhart and Rogoff 2008). Hence, in such periods, credit supply and demand may remain weak, irrespective of the interest rate set by the monetary authority, thus hampering the credit channel of monetary policy. The interest rate channel of monetary policy may be impaired, because in times of high uncertainty investors may postpone irreversible investment decisions until more information arrives (see, e.g., Bernanke 1983; Dixit and Pindyck 1994). Uncertainty then becomes a major determinant of investment decisions, whereas monetary policy loses its impact (Bloom, Bond, and Reenen 2007). Similarly, the interest rate responsiveness of investment may decline when firms and consumers have low confidence in their business or employment prospects (Morgan 1993). Finally, it may become more difficult for central banks to stabilize output because in times of high macroeconomic volatility firms tend to adjust their prices more often (Vavra 2014).

On the other hand, a monetary policy intervention could also be particularly effective, if it can mitigate some of the adverse financial crisis characteristics and thus prevent adverse feedback loops between the financial sector and the real economy, thereby restoring the functioning of the credit and interest rate channel (see, e.g., Mishkin 2009). In particular, credit constraints are more likely to bind during financial crises, leading to an increase in the external finance premium. Monetary policy may in this situation be able to decrease the external finance premium by easing credit constraints, so that the financial accelerator of Bernanke, Gertler, and Gilchrist (1999) would make monetary policy particularly effective. Moreover, while being less effective in the presence of high financial market distress and uncertainty, monetary policy can be all the more powerful if it is able to significantly alleviate financial market distress and reduce uncertainty (Bekaert, Hoerova, and Lo Duca 2013; Basu and Bundick 2017). Similarly, monetary policy can be more effective if it is able to raise confidence from very low levels, by providing signals about future economic prospects (Barsky and Sims 2012), by decreasing the probability of worst-case outcomes, and by improving the ability of agents to make probability assessments about future events (Ilut and Schneider 2014).[1]

It is therefore a priori ambiguous whether monetary policy is more or less effective during financial crises than during normal

[1]Bachmann and Sims (2012) provide evidence that the confidence channel is important for the effectiveness of fiscal policy in stimulating economic activity. Similar effects are conceivable in the context of monetary policy.

times. This ambiguity is further enhanced as crisis characteristics may vary over the course of a financial crisis (Borio and Hofmann 2017). Financial market distress and high uncertainty are primarily present in the initial and most acute phase of a financial crisis, when the economy is typically in a recession. In contrast, balance sheet adjustments may gain importance when financial market distress and uncertainty abate and the economy begins to recover.

Given the mixed theoretical predictions, the effectiveness of monetary policy during financial crises becomes an empirical question. However, such an analysis is made difficult, because financial crises are rare events and there is no unique definition of a financial crisis. We attempt to overcome these issues by exploiting the panel dimension of the data for twenty advanced economies over the period from 1984 to 2016 and by employing as measure of financial crises the widely used data set of Laeven and Valencia (2013), which identifies systematic banking crises via a transparent narrative approach. This approach captures systemic disruptions in the banking system, but it excludes periods during which financial stress is high for other reasons, such as political events or natural catastrophes. Our sample covers twenty financial crisis episodes that last 4.5 years on average. With fifteen out of twenty financial crisis episodes, our sample reflects to a large extent the global financial crisis that started in 2007, but also includes five earlier more regionally bounded crises.

We analyze differences in the transmission of monetary policy during financial crises and normal times using the interacted panel VAR (PVAR) methodology of Sá, Towbin, and Wieladek (2014), in which we interact the endogenous variables with a financial crisis indicator variable. This panel approach exploits the cross-section dimension of the data to improve the precision of the estimation, while allowing for a large degree of heterogeneity across countries. The PVAR model includes GDP and CPI as the main variables of interest and, in the baseline specification, the three-month interest rates as monetary policy instrument. We include several additional macroeconomic and financial variables, which are potentially important for the monetary policy transmission mechanism, such as credit, exchange rates, house prices, and stock market volatility as a proxy for uncertainty. We identify monetary policy shocks using a recursive identification.

In the baseline model, all variables in the VAR are interacted with financial crisis dummies in order to analyze differences in the transmission of monetary policy shocks between *financial crises* and *normal times*. In an extended model, we further investigate potential heterogeneity between different phases of financial crises: the *acute phase*, where the financial crisis is accompanied by a recession, and the *recovery phase*, where the financial crisis is ongoing but the economy is already in expansion. We use the Bry-Boschan algorithm to identify recession episodes, which is the most commonly used algorithm in the literature. The extended model can provide a more comprehensive understanding of monetary policy transmission during financial crises. However, it is also more prone to methodological limitations due to the reduced number of observations in each regime resulting from splitting financial crisis episodes into different phases, and due to the uncertainty inherent to the dating of recessions, on which our phase definition is built.

We find that the effects of monetary policy shocks on output and inflation are significantly larger and occur faster during financial crises than during normal times. Monetary policy shocks also explain a larger share of business cycle fluctuations during financial crises, according to forecast error variance decompositions. The effects of monetary policy shocks on uncertainty, credit, consumer confidence, and share prices are significantly larger compared with normal times, suggesting that these variables may play an important role in the transmission of monetary policy in financial crises. Further, we find that distinguishing between the acute and the recovery phases of financial crises can provide a more comprehensive picture of differences in the effects of monetary policy over the course of a crisis. An expansionary monetary policy shock has large and very quick positive effects on output and inflation during the acute phase of a financial crisis. On the other hand, during the recovery phase of a financial crisis the effects of monetary policy shocks occur at longer lags and have rather small effects on output and inflation, while the effects of financial variables are sizable. Finally, we find heterogeneity across countries with respect to the size and the timing of monetary policy actions during the global financial crisis. In the United States, monetary policy was eased quickly in 2008, mitigating the biggest drop in GDP at the end of 2008 and the beginning of 2009. By contrast, interest rates in the euro area were lowered about one year later, so that monetary policy shocks in many member states appear to have been contractionary in 2008 through the lens of our model and not to have mitigated the drop in GDP, as we show using counterfactual simulations. Euro-area monetary policy shocks became expansionary only at a later stage for most member states, in particular during the second recession in 2012/13, thus likely contributing to the final recovery.

We carefully check whether our results are affected by the zero lower bound and by unconventional monetary policies that have become relevant during the global financial crisis of 2008/09. In an alternative specification, we use estimated shadow interest rates from Wu and Xia (2016) to account for the zero lower bound and the effects of unconventional monetary policy for the United States, the United Kingdom, and the euro area. We also apply a more deterministic approach by excluding all observations in which the three-month interest rate is lower than 30 basis points.

Our finding that the effects of monetary policy are strong during financial crises and in particular during the acute phase, but less so in the recovery phase, are in line with findings of related empirical studies. Ciccarelli, Maddaloni, and Peydró (2013) analyze the effects of monetary policy in the euro area. They extend the end of their sample recursively from 2007 until 2011 and find, in line with our results, that monetary policy became more effective at stimulating economic activity during the global financial crisis, mainly via the credit channel. Dahlhaus (2017) focuses on the United States and finds that monetary policy is more effective in stimulating the economy during a high financial stress regime. By contrast, Bech, Gambacorta, and Kharroubi (2014) focus on the recovery period of financial crises by studying the effects of the monetary policy stance during recession episodes on the strength of the subsequent recovery for a panel of twenty-four advanced economies. They find that monetary policy is not very effective in stimulating GDP growth during the recovery phase of a financial crisis.[2]

[2]Other studies on monetary policy transmission during financial crises have focused on the effects of unconventional monetary policy (see, e.g., Gambacorta, Hofmann, and Peersman 2014 for a panel analysis or Williams 2014 for a survey). However, it is not possible to conclude whether monetary policy is more or less effective during financial crises compared with non-crisis times based on

The remainder of the paper is organized as follows. Section 2 describes the data set. Section 3 explains the econometric methodology. Section 4 presents and discusses the estimation results including various robustness checks. Finally, section 5 concludes.

2. Data

We first describe the panel data set of the endogenous variables for the PVAR and then our identification of financial crisis episodes, as well as the two phases of financial crises.

2.1 Data on Endogenous Variables

Our panel data set covers twenty advanced economies.[3] We use quarterly data for the period from the first quarter of 1984 to the fourth quarter of 2016, which provides us with a total of 2,640 observations.[4] Our baseline PVAR includes nine variables: real GDP, CPI, real house prices, the short-term interest rate, bank credit to the private non-financial sector, the effective exchange rate, consumer confidence, share prices, and stock price volatility.

GDP, CPI, three-month interest rates, effective exchange rates, and share prices data are taken from the OECD.[5] For the eleven countries in our sample that are members of the euro area, the aggregate euro-area three-month interest rates are used from 1999 onwards (from 2001 onwards in the case of Greece), accounting for their common monetary policy. Real house prices data come from

---

these studies because of the limited comparability of shocks to unconventional and conventional monetary policy measures.

[3] These are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States.

[4] The year 1984 was also chosen in previous studies that used panel models similar to ours (see, e.g., Sá, Towbin, and Wieladek 2014). Moreover, some of the endogenous variables, such as consumer confidence, are not available for several countries for longer samples. Starting our sample in 1970 instead of 1984 would only have added one financial crisis to our sample, i.e., the crisis in Spain between 1977 and 1981. Indeed, our results are robust when we estimate a model without consumer confidence from 1970 to 2016.

[5] Table A1 in the online appendix (available at http://www.ijcb.org) provides an overview of our data set and our data sources.

the International House Price Database from the Federal Reserve Bank of Dallas. Data for bank credit to the private non-financial sector are taken from the Bank for International Settlements. The consumer confidence measure is based on survey data and is obtained from various national sources via Thomson Financial Datastream. For example, for economies in the European Union, data stem from the Business and Consumer Surveys of the European Commission. Consumer confidence data are standardized to ensure an identical scale across countries. Finally, we follow Bloom (2009) and Cesa-Bianchi, Pesaran, and Rebucci (2018) and use stock market volatility as a measure of uncertainty. We calculate this measure as the average realized volatilities of daily stock returns for each quarter. The stock market indicators used represent the major stock indexes in each country and are obtained from Thomson Financial Datastream. We take logs of all variables except interest rates and consumer confidence indexes. Data not available in seasonally adjusted form are adjusted using a stable seasonal filter.

2.2 Financial Crises

We use the systematic banking crises identified by Laeven and Valencia (2013) to date financial crises. The data set is provided in annual frequency for the period from 1970 to 2011 for the twenty economies in our panel. Laeven and Valencia (2013) define the starting year of a banking crisis as when the two following criteria are met in the same year for the first time:

(i) Significant signs of financial distress in the banking system (as indicated by significant bank runs, losses in the banking system, and/or bank liquidations).

(ii) Significant banking policy intervention measures in response to significant losses in the banking system.

They consider the first criterion as a necessary and sufficient condition for a banking crisis. The second criterion is added as an indirect measure because financial distress is often difficult to quantify. They define the end of a banking crisis as the year before both real GDP growth and real credit growth remain positive for at least two consecutive years. In addition, they truncate financial crises after a maximum duration of five years.

We make two adjustments to the data set of Laeven and Valencia (2013) in order to make it applicable for our analysis. First, we transform their annual data set to a quarterly frequency. We follow the simplest and least restrictive approach and assume that a financial crisis begins in the first quarter of its first year and ends in the last quarter of its last year.[6] Second, the data set of Laeven and Valencia (2013) only covers up to 2011 and does not include the end of some of the financial crises that started in 2008. Therefore, we extend the database until the year 2016. We use the same criteria as Laeven and Valencia (2013) to define the end of financial crises, based on real GDP and credit data and a maximum duration of financial crises of five years.[7]

On the basis of this data set, we construct a financial crisis dummy $D_{i,t}$, which takes a value of one when there is a financial crisis in country i at period t, and a value of zero otherwise. In our main analysis, we distinguish between two regimes on the basis of this dummy: financial crises and normal times.

In an extended analysis, we further distinguish between an acute and a recovery phase of a financial crisis. We understand the acute phase of a financial crisis as a period during which the immediate impact of the crisis on the economy is most noticeable, e.g., via high volatility and weak economic activity. The recovery phase is understood as the period when this immediate impact of the crisis has largely disappeared but the legacy of the crisis may still impact the economy, e.g., via balance sheet adjustments. As there is no single indicator available that allows for a sharp distinction

[6] As shown in the online appendix, results are very similar when using a more restrictive approach for the transformation to the quarterly frequency, where financial crises are set to begin in the fourth quarter of their first year and end in the first quarter of the last year.

[7] We make one exception to the truncation scheme of Laeven and Valencia (2013). In seven countries in our sample, the recession which began during the financial crisis in 2008 was still ongoing for some quarters in 2013. In these countries, we assume that the financial crisis ends together with the recession in 2013 instead of truncating it at the end of 2012. This extension only affects twelve quarterly observations in total and does not drive our results, but it assures a more consistent treatment of recessions that arise during financial crises, as opposed to normal recessions.

Table 1. Summary Statistics: Financial Crisis and Recession Episodes

	Quarters	Share	Episodes	Average Length (Quarters)
Financial Crises				
Total	367	14%	20	18
Acute Phase (FC + Recession)	190	7%	30	6
Recovery Phase (FC + Expansion)	177	7%	34	5
Normal Times				
Total	2,273	86%		
Recession	196	7%	56	4
Expansion	2,077	79%		

Note: The table shows data for 2,640 total observations over twenty countries between 1984:Q1 and 2016:Q4.

between these two phases, we follow the literature on asymmetric effects of monetary policy over the business cycle and use business cycle phases to distinguish between acute phases and recovery phases of financial crises. For this purpose, we use the Harding and Pagan (2002) version of the Bry-Boschan algorithm, which is the most frequently applied business cycle dating algorithm in the literature, to define—similarly to the financial crisis dummy indicator—a recession dummy.[8] Based on the financial crisis dummy and the recession dummy, we define four regimes: the acute phase of a financial crisis (financial crisis and recession), the recovery phase of a financial crisis (financial crisis and expansion), and recessions and expansions in normal times.

Table 1 shows some summary statistics regarding the number of observations, and the length and frequency of financial crisis

[8]This algorithm identifies local peaks and troughs in real GDP and defines a recession as the period from the peak to the trough and an expansion as the period from trough to peak.

episodes. There are twenty financial crisis episodes in our sample. With an average length of eighteen quarters, financial crises are rather persistent. The episodes cover 367 quarters in the sample and thus about 14 percent of the total 2,640 observations. In addition, there are eighty-six recession episodes in the sample, of which thirty occur during financial crises and fifty-six during normal times. Recession episodes are much less persistent than financial crises, with an average length of six quarters when they occur during a financial crisis and a length of four quarters in normal times. Within financial crisis episodes, there is a roughly similar number of acute and recovery episodes of a length of five to six quarters, as about half of the financial crisis quarters are recessionary. All twenty financial crisis episodes are accompanied by at least one recession, which usually occurs at the beginning of a financial crisis. However, in some cases—in particular, during the global financial crisis that started in 2007—two separate recessions occurred during the financial crisis, with a recovery phase in between.

In terms of the regional distribution, two financial crises occurred in Sweden and the United States, no financial crises occurred in Australia and Canada, whereas each of the other economies experienced one financial crisis. The number of recessions per economy ranges from two (Canada, Australia, Netherlands, and Ireland) to seven (Greece). Our sample of financial crises is dominated by the global financial crisis that started in 2007, as fifteen out of twenty financial crises took place in this period. Apart from that, one financial crisis occurred in the United States during 1988; three during the early 1990s in Finland, Norway, and Sweden; and one in Japan around the year 2000.

3. Methodology

We base our empirical model on the Bayesian interacted PVAR methodology developed in Sá, Towbin, and Wieladek (2014), which extends the model of Towbin and Weber (2013) using cross-country heterogeneous coefficients. The model exploits the cross-country dimension of the data, accounts for the dynamics between the main macroeconomic variables, and allows for interactions between macroeconomic variables and exogenous terms, i.e., financial crisis and recession dummy variables in our study.

3.1 Interacted PVAR Model

We start by describing the simple PVAR model without interaction terms.[9] The PVAR in structural form is given by

$$J_i Y_{i,t} = \tilde{A}_{i,0} + \sum_{k=1}^{L} \tilde{A}_{i,k} Y_{i,t-k} + \tilde{u}_{i,t}, \tag{1}$$

where $t = 1, \ldots, T$ denotes time, $i = 1, \ldots, N$ denotes the country, $Y_{i,t}$ is a $q \times 1$ vector of endogenous variables, $\tilde{A}_{i,0}$ is a vector of country-specific intercepts, and $\tilde{A}_{i,k}$ is a $q \times q$ matrix of autoregressive coefficients up to lag L.[10] The $q \times q$ matrix of contemporaneous effects J_i is lower triangular with a vector of ones on the main diagonal. The $q \times 1$ vector of structural residuals $\tilde{u}_{i,t}$ is assumed to be normally distributed with a mean of zero and with a diagonal $q \times q$ covariance matrix $\tilde{\Sigma}$.

We follow Sá, Towbin, and Wieladek (2014) by allowing for heterogeneous intercepts $\tilde{A}_{i,0}$ and heterogeneous slope parameters $\tilde{A}_{i,k}$ across countries. Moreover, we also allow the parameters of the contemporaneous coefficients matrix J_i to vary across countries by estimating the model recursively, i.e., directly in its structural form. Country-specific coefficients result from adding dummy variables for each country into the model. In order to compute the average effect over countries, we assign the value of $1/N$ to each of the country-specific dummies. For the PVAR without interaction terms, this approach is equivalent to the mean group estimator of Pesaran and

---

[9]PVARs have been used in various empirical analyses of monetary policy to improve the precision of the estimation and to detect common country dynamics. For example, see Assenmacher-Wesche and Gerlach (2008), Goodhart and Hofmann (2008), Carstensen, Hülsewig, and Wollmershäuser (2009), Calza, Monacelli, and Stracca (2013), and Gambacorta, Hofmann, and Peersman (2014).

[10]We assume that there are no dynamic interdependencies across countries, i.e., the endogenous variables of country i are not affected by other countries' variables and residuals are uncorrelated across countries. This can be a rather strong assumption in a cross-country framework, but it drastically reduces the number of parameters that need to be estimated, and thus it is used frequently in PVAR applications.

Smith (1995) and it gives consistent estimates for dynamic panels in the presence of cross-country heterogeneity.[11]

The PVAR is then augmented with interactions between the endogenous variables and an exogenous indicator variable. The structural interacted PVAR (IPVAR) model with a single interaction term is given by

$$J_{i,t}Y_{i,t} = \tilde{A}_{i,0} + \sum_{k=1}^{L} \tilde{A}_{i,k}Y_{i,t-k} + \tilde{B}_0 D_{i,t} + \sum_{k=1}^{L} \tilde{B}_k D_{i,t} Y_{i,t-k} + \tilde{u}_{i,t}, \tag{2}$$

$$J_{i,t}(w,q) = J_i(w,q) + J(w,q)D_{i,t}, \tag{3}$$

where $\tilde{A}_{i,0}$, $\tilde{A}_{i,k}$, and $\tilde{u}_{i,t}$ are defined as in equation (1). $D_{i,t}$ is the financial crisis dummy variable, which takes a value of one if there is a financial crisis in country i during period t, and a value of zero otherwise. $\tilde{B}_0$ are intercepts and $\tilde{B}_k$ is a $q \times q$ matrix of autoregressive coefficients up to lag L, which correspond to the financial crisis regime, respectively. The matrix of contemporaneous coefficients $J_{i,t}$ is again lower triangular with a vector of ones on the main diagonal.

The coefficients that capture the dynamics in normal times, $\tilde{A}_{i,0}$ and $\tilde{A}_{i,k}$, remain country specific through the addition of country dummies. By contrast, the coefficients that capture changes in the dynamics during financial crises, $\tilde{B}_0$ and $\tilde{B}_k$, are assumed to be homogeneous across countries. We make a similar assumption for the contemporaneous coefficients $J_{i,t}$. Equation (3) shows that each lower triangular element (w,q) of $J_{i,t}$—where w indicates the row, q indicates the column, and $q < w$—comprises a country-specific part, $J_i(w,q)$, for the non-crisis regime and a regime-specific part, $J(w,q)D_{i,t}$, which does not vary across countries.

[11]Pesaran and Smith (1995) estimate separate VAR models for each country and then calculate averages over the estimated coefficients, whereas Sá, Towbin, and Wieladek (2014) introduce cross-country heterogeneity by augmenting parameters with country dummies. The latter approach allows the combination of cross-country heterogeneous coefficients and country-invariant interaction terms in the interacted PVAR model.

Pre-multiplying the recursive-form IPVAR in equation (2) with $J_{i,t}^{-1}$ yields the reduced-form version of the model:

$$Y_{i,t} = A_{i,0} + \sum_{k=1}^{L} A_{i,k} Y_{i,t-k} + B_{i,0} D_{i,t} + \sum_{k=1}^{L} B_{i,k} D_{i,t} Y_{i,t-k} + u_{i,t}, \tag{4}$$

with the reduced-form coefficient matrices defined as, e.g., $A_{i,0} = J_{i,t}^{-1} \tilde{A}_{i,0}$. Note that the reduced-form coefficients corresponding to the financial crisis regime, $B_{i,0} = J_{i,t}^{-1} \tilde{B}_0$ and $B_{i,k} = J_{i,t}^{-1} \tilde{B}_k$, as well as the residuals $u_{i,t} = J_{i,t}^{-1} \tilde{u}_{i,t}$ and their covariance matrix $\Sigma_{i,t} = J_{i,t}^{-1} \tilde{\Sigma} (J_{i,t}^{-1})'$, will also vary across countries and regimes due to the variation in $J_{i,t}$. Thus, estimating the model recursively across different regimes allows for heteroskedasticity of residuals and for country- and regime-specific contemporaneous correlations.

Using the above approach, we exploit the panel structure of our data and address the problem of the small number of financial crisis observations per country. In addition, we allow for as much cross-country heterogeneity as possible and we assume that only the recursive-form coefficients that capture the dynamics during financial crises are homogeneous across countries. By contrast, the reduced-form coefficients in financial crises vary across countries via country-specific contemporaneous effects.[12] In this, our approach allows for more cross-country heterogeneity than a fixed-effect estimator does and therefore provides estimates that are less likely to suffer from inconsistency.[13] At the same time, the use of country

[12]Georgiadis (2014) shows that the heterogeneity of monetary policy transmission is quite large across advanced economies and it can be explained mainly by differences in the financial structure, labor market rigidities, and the industry mix.

[13]When we estimate the IPVAR with fixed effects (i.e., only allowing for country-specific intercepts), the impulse responses to a monetary policy shock are highly persistent and the responses of CPI exhibit signs of the price puzzle. This may be attributable to the fixed-effects estimator being biased in dynamic panels when the time dimension is small (Nickell 1981) as well as when the time dimension is large, but the coefficients of lagged endogenous variables are heterogeneous across countries (Pesaran and Smith 1995). The pooled mean group estimator used by Pesaran and Smith (1995) is only biased if the time dimension is small. In our case, $T = 120$, so the bias from this source should be

dummies to allow for cross-country heterogeneity is more flexible than the mean group estimator of Pesaran and Smith (1995). It allows us to combine country-specific and pooled effects within a PVAR when facing a tradeoff between heterogeneity in the data and the aim of pooling information across countries in order to increase the precision of estimated parameters.

The impulse responses of endogenous variables to a monetary policy shock can be computed from the reduced form of the VAR and the identified structural monetary policy shocks. In the IPVAR model, the impulse responses depend on the values of the interaction terms, and thus on the regime. We assume that the economies remain in the regime that prevailed when the monetary policy shock occurred. Thereby, we follow most previous studies that analyze asymmetries in monetary policy transmission during recessions and expansions and during periods of high financial stress using VAR methods.[14]

In the second part of the analysis, we extend model (2) to distinguish between four regimes. For this purpose, we define $D_{i,t}$ not as a single dummy variable but as a vector of three dummy variables, $D_{i,t} = [D_{i,t}^{FC}, D_{i,t}^{R}, D_{i,t}^{FCR}]$, where $D_{i,t}^{FC}$ is the financial crisis dummy, $D_{i,t}^{R}$ is the recession dummy, and $D_{i,t}^{FCR} = D_{i,t}^{FC} * D_{i,t}^{R}$ is the interaction term between the first two dummies, taking a value of one when there is a financial crisis *and* a recession in country i at time t, and zero otherwise. These dummies allow us two distinguish between the following regimes:

- $D_{i,t}^{FC} = 1$, $D_{i,t}^{R} = 1$, $D_{i,t}^{FCR} = 1$: acute phase of a financial crisis (crisis and recession),

negligible. The introduction of interaction terms that are assumed to be homogeneous across countries introduces some bias to the extent that the homogeneity assumption is violated by the data. However, we consider our setup as a good solution to the tradeoff between minimizing bias and enabling the estimation of financial crisis effects.

[14]An exception is Fry-McKibbin and Zheng (2016), who calculate regime-independent, non-linear impulse responses in a threshold VAR model for the United States. They find that only large monetary policy shocks of two standard deviations or more are able to endogenously move the economy from a low to a high financial stress regime or vice versa, while small shocks do not increase the probability of regime switches. Hence, imposing the regimes exogenously rather than allowing for regime switches in response to monetary policy shocks does not seem to be a very restrictive assumption.

- $D_{i,t}^{FC} = 1$, $D_{i,t}^{R} = 0$, $D_{i,t}^{FCR} = 0$: recovery phase of a financial crisis (crisis, no recession),
- $D_{i,t}^{FC} = 0$, $D_{i,t}^{R} = 1$, $D_{i,t}^{FCR} = 0$: normal recession (no crisis, recession),
- $D_{i,t}^{FC} = 0$, $D_{i,t}^{R} = 0$, $D_{i,t}^{FCR} = 0$: normal expansion (no crisis, expansion).

3.2 Estimation and Inference

We follow Uhlig (2005) and Sá, Towbin, and Wieladek (2014) and estimate the model with Bayesian methods. As argued by Uhlig (2005), the Bayesian approach has the advantage that it is computationally simple and that it allows us to draw error bands for impulse responses and variance decompositions using Monte Carlo simulations. At the same time we use a flat prior, which does not impose any shrinkage on the parameters. The only additional prior information we incorporate is that of a stable VAR, as we discard explosive draws.

In particular, we apply a conjugate normal-Wishart prior which allows us to obtain joint draws from the posterior distribution by direct Monte Carlo sampling of all recursive-form parameters (Uhlig 1994, 2005; Koop and Korobilis 2010; Del Negro and Schorfheide 2011). The normal-Wishart distribution specifies that Σ^{-1} follows a Wishart distribution $\mathcal{W}_m(S^{-1}/\nu, \nu)$, where S is the mean covariance matrix and ν are the degrees of freedom describing the uncertainty about $(\mathbf{B}, \Sigma)$ around the mean coefficients $(\bar{B}, S)$. Conditionally on Σ, the matrix of vectorized recursive coefficients $vec(\mathbf{B})$ follows a normal distribution $\mathcal{N}(vec(\bar{B}), \Sigma \otimes K^{-1})$, where $\bar{B}$ is an $l \times m$ dimensional matrix of mean coefficients and K is an $l \times l$ positive definite matrix.

The normal-Wishart posterior has a closed-form solution. As we use a flat prior, the posterior mean of the normal distribution corresponds to the maximum-likelihood estimator of the recursive coefficient matrix $\hat{B}$, and the posterior covariance matrix corresponds to $\hat{S}$, i.e., the sum of squared residuals from the maximum-likelihood estimation of equation (2). The number of degrees of freedom is set to equal the number of observations over time and countries (excluding missing values). We choose two lags in the baseline IPVAR specification according to the Akaike information criterion.[15]

For each draw from the posterior, we evaluate the interaction terms $D_{i,t}$ at their values of interest. We then compute the reduced-form parameters according to equation (4) by pre-multiplying the structural parameters with the inverted lower-diagonal matrix of contemporaneous coefficients, $J_{i,t}^{-1}$, and we calculate the impulse responses for each regime. Here, we impose the additional prior that responses are non-explosive, as in Uhlig (1994) and Cogley and Sargent (2005), i.e., we discard draws that lead to explosive parameters. We continue drawing from the posterior until we have 1,000 non-explosive draws, which we use to compute the median and 90 percent probability bands.

3.3 Identification of Monetary Policy Shocks

We estimate the interacted PVAR in its recursive form. Hence, monetary policy shocks are identified implicitly by the ordering of the variables in the VAR. We order GDP, CPI, and house prices before the short-term interest rate and we assume that these variables do not react to a monetary policy shock on impact. This assumption has been employed in a large number of VAR studies and agrees with the theoretical predictions from DSGE models, i.e., a delayed response of output and prices to monetary policy. We order banks' credit to the private sector, the exchange rate, consumer confidence, share prices, and stock market volatility after the interest rate, and thus allow these variables to react to monetary policy shocks on impact. This is consistent with the view that financial market variables or variables of market confidence are less rigid, so they can react to monetary policy shocks contemporaneously. Nonetheless, our results are robust when we choose alternative orderings, such as ordering house prices after the interest rate or ordering banks' credit to the private sector and consumer confidence before the interest rate.

As a robustness check, we also identify monetary policy shocks using the sign-restrictions approach developed in Canova and De Nicoló (2002) and Uhlig (2005), which was applied in a PVAR

[15] Two lags are also used in comparable PVAR models; see Calza, Monacelli, and Stracca (2013) and Sá, Towbin, and Wieladek (2014).

context in Carstensen, Hülsewig, and Wollmerhäuser (2009), Gambacorta, Hofmann, and Peersman (2014), and Sá, Towbin, and Wieladek (2014). This approach has the advantage that identification is independent of the ordering in the VAR and no zero contemporaneous restrictions need to be imposed. Instead, the sign of the impulse responses of a subset of endogenous variables is restricted in line with theoretical predictions. However, the latter can also be restrictive, particularly in the analysis of financial crises, when the economy might behave very differently from the predictions of standard theoretical models. We only restrict the impulse response of the short-term interest rate to be negative during the first four quarters and the response of CPI to being positive between the second and fourth quarter, whereas we remain agnostic about the responses of GDP and the other variables in the VAR.

4. Results

In this section, we describe our results for the baseline model that differentiates between financial crises and normal times, followed by the results of the extended model, where we distinguish between the acute and the recovery phases of financial crises. Subsequently, we discuss our results in light of related empirical and theoretical findings of the literature, and we provide some robustness checks.

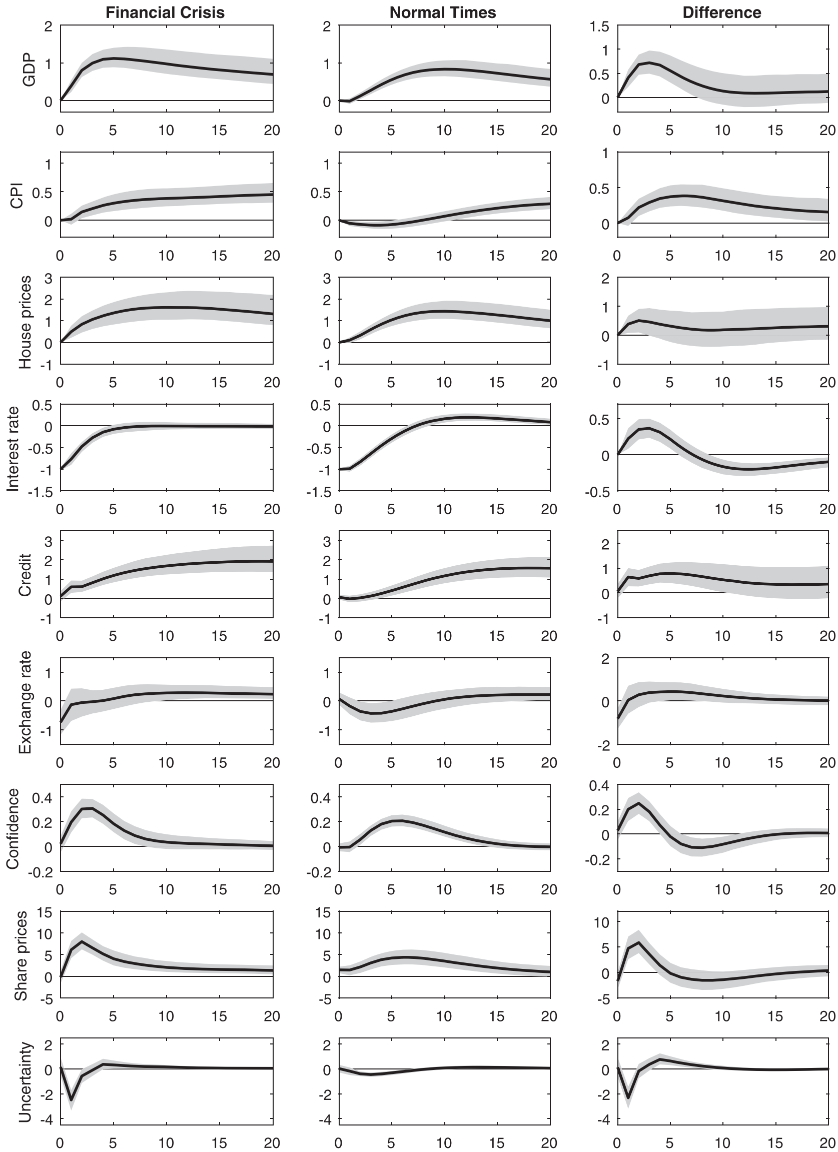
4.1 Monetary Policy Transmission in Financial Crisis and Normal Times

We analyze differences in monetary policy transmission during financial crises and normal times based on an impulse response analysis, a forecast error variance decomposition, plots of identified monetary policy shocks, and a counterfactual scenario that shows trajectories of macroeconomic variables during financial crises based on the assumption of a constant interest rate.

4.1.1 Panel Impulse Response Analysis

Figure 1 shows the impulse responses and the 90 percent probability bands to a 1 percentage point expansionary monetary policy shock in the baseline model. The first column shows the effects of

Figure 1. Effects of a Monetary Policy Shock during Financial Crises and Normal Times



Notes: Median impulse responses to a 100 basis point expansionary monetary policy shock together with 90 percent probability bands, in percent. For each draw from the posterior, impulse responses are computed for the different regimes. Column 3 shows median differences between the two regimes and the corresponding probability bands in percentage points.

a monetary policy shock during financial crises, the second column shows the effects during normal times, and the third column shows the difference between the two. All impulse responses correspond to average effects across all twenty countries in our sample.

During normal times, our results for output and prices are in line with the results of previous empirical studies (see, e.g., Christiano, Eichenbaum, and Evans 1999). The response of output is hump shaped and is faster than the response of prices, which start increasing after about two years. Monetary policy shocks have permanent effects on the price level, whereas the effect on output returns to zero in the long run.

However, during financial crises the responses of output and prices are significantly stronger and occur faster. Within the first two years after an expansionary monetary policy shock, the response of output is up to 0.8 percentage point higher compared with normal times; subsequently, there is no significant difference between the two regimes. Prices increase within the first few quarters after the shock during financial crises. The differences in the responses between the regimes are rather persistent: over the entire horizon, the increase in prices is about 0.3 percentage point higher than that during normal times. The impulse responses of the interest rate shows that the larger effects of monetary policy shocks on output and prices during financial crises are not driven by differences in the systematic part of monetary policy: in financial crises, an expansionary monetary policy shock has a less persistent effect on the interest rate during the first two years.

We also find significant differences between the two regimes for most of the other variables. After a monetary policy shock, credit increases more instantaneously and, for about two years, significantly stronger during financial crises. Similar patterns are observed for the other variables, but differences between the responses in the two regimes are rather short-lived, lasting for up to one year. Uncertainty exhibits a marked decline of up to 2.5 percent after an expansionary monetary policy shock during financial crises, whereas it barely reacts during normal times. Share prices increase by up to 6 percentage points stronger in financial crises. Consumer confidence also increases stronger, but the effect returns to zero faster than in normal times. House prices and the exchange rate react slightly stronger during financial crises for the first two quarters after

Table 2. Forecast Error Variance Explained by the Monetary Policy Shock, Two Regimes

	Financial Crisis		
	One Year	**Two Years**	**Five Years**
GDP	22.9 (14.6, 32.6)	35.1 (24.9, 46.5)	39.9 (30.6, 53.8)
CPI	6.6 (2.7, 13.3)	15.7 (7.6, 26.7)	33.4 (21.0, 47.2)
House Prices	7.3 (3.5, 14.0)	18.1 (9.5, 28.5)	29.5 (16.9, 42.1)
Interest Rates	70.8 (61.2, 78.0)	59.6 (51.2, 69.6)	52.7 (44.2, 63.6)
Credit	11.2 (4.1, 19.9)	24.1 (11.8, 33.7)	42.4 (29.6, 55.5)
Exchange Rate	2.3 (0.7, 6.5)	2.7 (0.9, 8.5)	4.0 (1.7, 11.2)
Confidence	22.0 (15.1, 30.3)	22.9 (15.9, 29.5)	21.8 (15.0, 28.6)
Share Prices	25.3 (15.8, 34.2)	27.9 (18.8, 38.6)	28.4 (20.0, 39.1)
Uncertainty	14.0 (8.1, 22.2)	13.9 (7.9, 22.5)	14.1 (8.3, 22.5)
	Normal Times		
	One Year	**Two Years**	**Five Years**
GDP	4.1 (1.5, 8.2)	16.0 (7.6, 26.1)	28.0 (14.6, 39.0)
CPI	0.4 (0.0, 2.4)	0.8 (0.3, 3.0)	11.1 (4.8, 21.5)
House Prices	5.5 (2.3, 9.3)	15.4 (7.9, 24.1)	26.4 (14.7, 37.2)
Interest Rates	85.7 (80.1, 89.9)	74.8 (65.8, 81.0)	69.1 (60.0, 74.8)
Credit	1.1 (0.3, 2.7)	7.8 (2.0, 14.5)	25.8 (13.1, 37.2)
Exchange Rate	1.0 (0.1, 3.9)	1.8 (0.2, 7.8)	2.9 (0.9, 8.7)
Confidence	9.2 (5.2, 15.4)	20.6 (12.8, 30.6)	21.6 (13.0, 32.5)
Share Prices	2.9 (0.7, 7.7)	7.8 (2.4, 17.9)	12.2 (3.9, 24.5)
Uncertainty	2.5 (0.9, 5.3)	3.1 (1.4, 6.8)	3.6 (1.7, 7.9)

Notes: Median share of the forecast error variance explained by the monetary policy shock, in percent. The sixteenth and eighty-fourth credibility sets are shown in parentheses.

the shock, but differences between the regimes turn insignificant thereafter.

Our main finding that the effects of monetary policy shocks are larger during financial crises is reflected in the forecast error variance decompositions shown in table 2. Our results for normal times are in line with the common finding in the literature that monetary policy shocks explain only relatively small fractions of forecast errors of output and inflation (see, e.g., Bernanke, Boivin, and Eliasz 2005; Coibion 2012). However, during financial crises this fraction is considerably higher for output and prices, and the differences between the regimes are particularly pronounced at low horizons. For example, the monetary policy shock explains about 4 percent of the forecast error of output at the one-year horizon during normal times, but more than 20 percent in the financial crisis regime. For prices, the shares of forecast errors explained by monetary policy shocks in normal times are still close to zero at the one- and two-year horizons and only increase at higher horizons, while during financial crises the share is at above 6 percent at the one-year horizon. Also for all other variables—most notably credit, share prices, uncertainty, and confidence—the monetary policy shock explains considerably larger fractions of the total forecast error, particularly at low horizons, in the financial crisis regime compared with normal times.

4.1.2 Cross-Country Heterogeneity

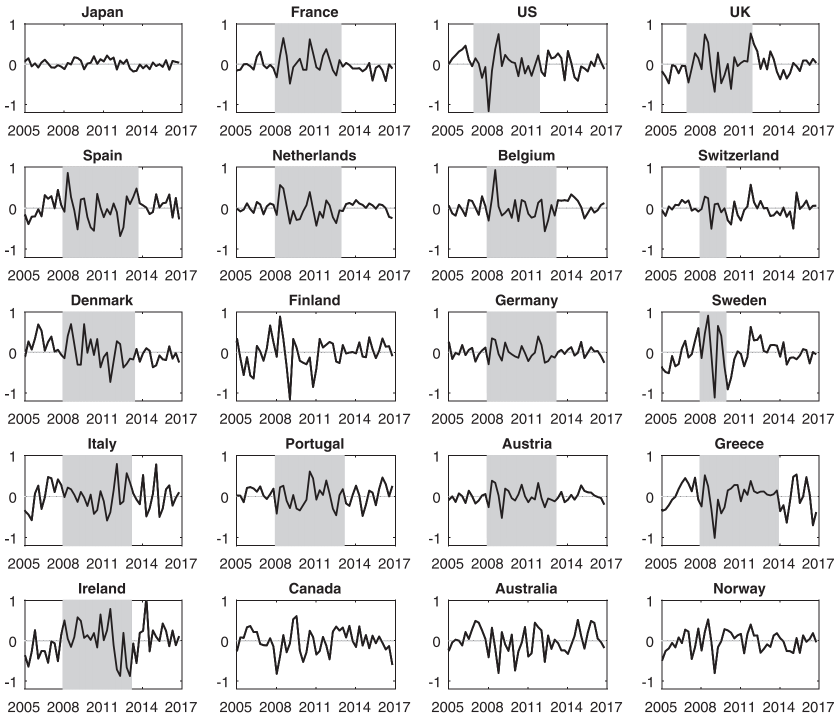
The results presented until now refer to the average effects across twenty economies. However, to the extent that monetary policy reacted differently during financial crisis episodes in individual countries—for instance, by intervening more or less aggressively or at different points in time—cross-country heterogeneity can be important and potentially helpful for the interpretation of the results. We therefore show the series of the identified monetary policy shocks for individual countries in order to assess whether and to what extent monetary policy surprises were expansionary or contractionary during financial crisis episodes. Further, for selected countries, we compute counterfactual paths of the endogenous variables under the assumption of a constant interest rate over financial crisis episodes and compare them with actually observed paths.[16]

Figure 2 shows the identified monetary policy shocks for each of the twenty countries in the sample.[17] We restrict the plots to the

[16]In principle, the PVAR model also allows us to compute impulse responses on a country-by-country basis, as the reduced-form coefficients in the model vary across countries. However, such results are subject to instability and a high degree of estimation uncertainty due to the very low number of observations per country during financial crises.

[17]While all euro-area member states in the sample were subject to a common monetary policy and thus similar short-term interest rate paths since 1999, the PVAR model estimates separate interest equations for each country. We can interpret the identified monetary policy shocks in the individual euro-area countries

Figure 2. Estimated Monetary Policy Shocks: 2005–17



Notes: Residuals corresponding to the interest rate equation of each country from the structural form of the IPVAR of equation (2). Gray bars show the financial crisis episodes for each country.

2005–17 subperiod to focus on cross-country differences during the global financial crisis and its aftermath. The shock series are by their nature quite volatile, typically exhibiting alternating expansionary and contractionary movements. They add up to zero for each country over the total sample period, but can deviate from zero during subperiods, and thus also during financial crisis episodes.

---

as country-specific monetary policy surprises, given the economic conditions of each country according to the model. This is in line with Barigozzi, Conti, and Luciani (2014) and Corsetti, Duarte, and Mann (2018), who show that euro-area monetary policy shocks transmit heterogeneously across member states.

The United States exhibits the largest expansionary shock directly at the beginning of the financial crisis, and it took some time until that expansionary movement was followed by a contraction, which was likely related to the zero lower bound. The United Kingdom shows a similar pattern, although the respective shocks were much smaller. By contrast, monetary policy shocks were contractionary for most euro-area member states at the onset of the global financial crisis, reflecting the delayed reduction in common euro-area interest rates compared with the United States. Given the large effects of monetary policy shocks during financial crises documented above, these differences in the direction of monetary policy surprises may have partially contributed to the differences in macroeconomic performance between the United States, the United Kingdom, and the euro area. While the United States was the epicenter of the global financial crisis, there was only one recession in 2008 and 2009 associated with this crisis. The global financial crisis spread to Europe, causing a recession of similar length in the euro area and the United Kingdom as in the United States, but then the euro area fell into a longer second recession related to the European sovereign debt crisis in 2011 that lasted for two years, while there was no such second recession in the United States and the United Kingdom. During that time, the monetary policy surprises became expansionary for most euro-area member economies, contributing to the final recovery.

Finally, we complement our analysis with a counterfactual exercise based on a constant interest rate scenario.[18] We compute a series of consecutive one-step-ahead forecasts on a country-by-country basis for the main macroeconomic variables in our model for specific financial crisis episodes, conditional on a predetermined counterfactual constant interest rate path for the short-term interest rate using the approach by Lütkepohl (2013).[19]

[18]Chen et al. (2016) apply similar counterfactual forecasts to evaluate the global effects of U.S. quantitative easing measures within a global vector autoregressive (GVAR) model. Giannone et al. (2012) use a similar counterfactual exercise to gauge the effects of non-standard monetary policy measures in the euro area by conditioning on different paths of interest rate spreads.

[19]We evaluate the PVAR model for each country individually. For each financial crisis episode in a particular country, we define the time period t^* as the quarter when the episode started. From that point in time onwards, we restrict

By comparing the counterfactual paths with the actual data, we can assess the effects of interest rate adjustments during financial crisis episodes. To the extent that monetary policy is in control of the short-term interest rate and that our model reflects the monetary policy reaction function of central banks, the counterfactual analysis provides information on changes in the interest rate due to the systematic monetary policy of central banks (according to their reaction function) as well as due to monetary policy shocks. However, the Lucas critique implies that endogenous relations between macroeconomic variables might change relative to our baseline model, as the interest rate path is now fixed instead of reacting endogenously. We thus interpret the results of this counterfactual exercise with caution.

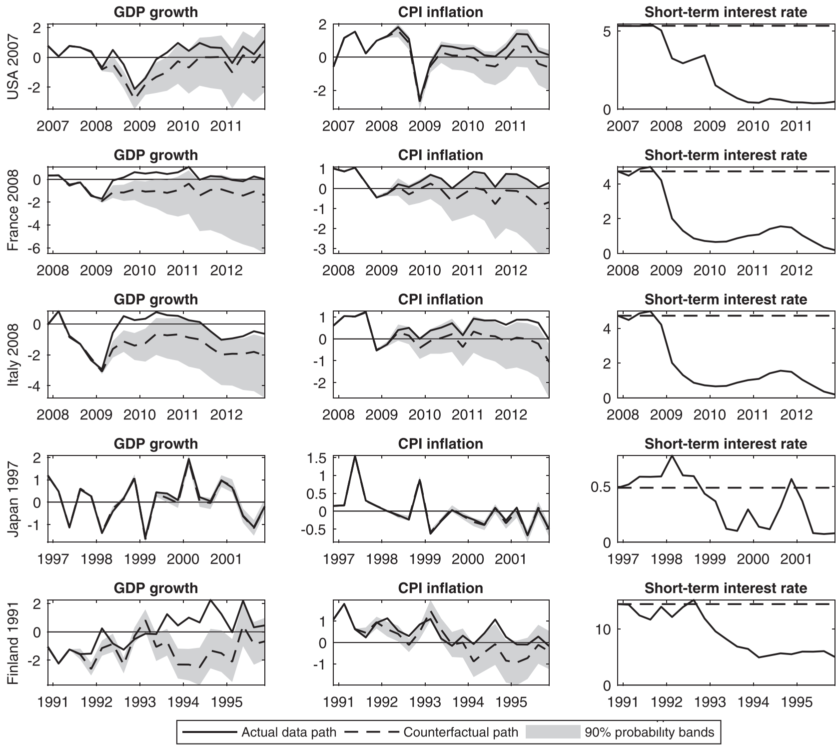
Because a discussion of the results for all countries in our sample is beyond the scope of this exercise, we focus on five countries. We look at the United States and two euro-area economies (France and Italy) during the global financial crisis and also study two earlier more regionally bounded crises in Japan (1997–2001) and Finland (1991–95). Figure 3 presents the counterfactual paths (dashed lines) of the main variables for the years of the financial crisis episode in the respective country and 90 percent probability bands, together with the actual outcomes (solid lines) of the respective variables.

For the United States, we find that GDP growth and inflation are considerably lower in the counterfactual scenario of no change in the interest rate (interest rate remains at about 5 percent) compared with the actual paths. GDP growth (quarter-on-quarter) is 1 to 2 percentage points lower in this scenario from 2008 onwards, and inflation remains in negative territory for most of this period. The differences between the actual and the counterfactual paths are even larger in France and, in particular, in Italy from mid-2009 onwards. Hence, expansionary monetary policy, given its large effects during

---

the short-term interest rate of the respective country to remain unchanged at the level it assumed at time $t^* - 1$, i.e., $\{int_{i,t}\}_{t=t*}^{t=T} = int_{i,t^*-1}$, while all other variables in the VAR are endogenous. We then compute one-step-ahead forecasts of the endogenous variables $\tilde{y}_{i,t}^{t^*}$, given their values up to time $t^* - 1$ and the lagged value of the short-term interest rate. Further one-step-ahead conditional forecasts $\tilde{y}_{i,t}^{t^*+1}$ up to $\tilde{y}_{i,t}^{t^*+20}$ are computed on the basis of the previous forecasts and the exogenous interest rate path.

Figure 3. Counterfactual Paths for Selected Financial Crisis Episodes



Notes: Black solid lines show actual paths of the series during selected financial crisis episodes in the United States (financial crisis 2007–11), France and Italy (2008–12), Japan (1997–2001), and Finland (1991–95). Dashed lines show counterfactual paths based on consecutive one-step-ahead forecasts from the interacted PVAR model, starting from the first quarter of the financial crisis ($t = t^*$), where the short-term interest rate is held fixed for $t \geq t^*$, together with 90 percent probability bands.

financial crises as indicated by our model, may have helped to prevent a much longer recession. At the same time, differences in timing might have mattered. In the United States, the interest rate already decreased considerably in 2008, mitigating the adverse effects of the crisis on GDP during that time. By contrast, in France and Italy the decrease in the interest rate started only at the end of 2008, and the downturn of GDP in 2008 is therefore equally pronounced in the counterfactual scenario as in the actual data. An earlier cut in interest rates could have possibly mitigated the large drop in GDP in 2008 and the beginning of 2009 in particular, given that monetary policy transmits quickly and strongly to the economy according to our results.

The large effects of monetary policy hold not only for the recent global financial crisis but also for earlier financial crisis episodes, as shown for the example of Finland. On the other hand, for the Japanese crisis starting in 1997, there are basically no differences in GDP growth and inflation between the counterfactual and the actual data paths, most likely because the short-term interest rate was already close to zero at the beginning of the crisis, so there was little room for further interest rate cuts.

4.2 Acute and Recovery Phases of Financial Crises

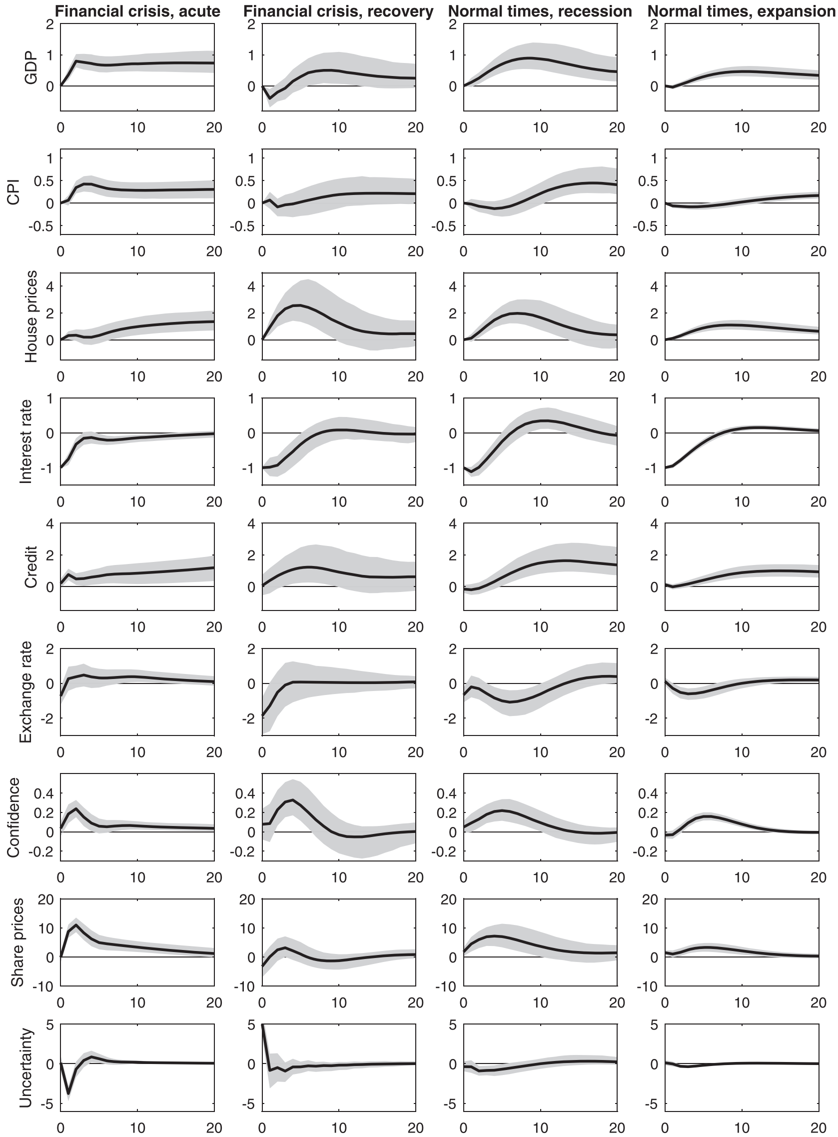
Figure 4 shows the impulse responses to an expansionary monetary policy shock in the model that differentiates between four regimes: acute phase of financial crises, recovery phase of financial crises, recessions in normal times, and expansions in normal times. The effects of monetary policy shocks differ between the different regimes, particularly between the acute and recovery phases of financial crises.

The effects of an expansionary monetary policy shock during the acute phase of a financial crisis are roughly similar to the results for financial crises in the model with two regimes: an expansionary monetary policy shock leads to a rapid and persistent increase in output and prices. Also, share prices increase while uncertainty decreases strongly in this regime.

By contrast, during the recovery phase of a financial crisis, a monetary policy shock has significantly lower and more-delayed effects on output and prices.[20] The effect on output becomes significantly positive only after about two years, according to the 90 percent probability bands. The response of prices becomes positive after two years, but is not significantly different from zero. Share prices

[20]Figure A2 in the online appendix illustrates the differences in the impulse response functions.

Figure 4. Effects of Monetary Policy Shocks in Four Regimes



Notes: Median impulse responses to a 100 basis point expansionary monetary policy shock together with 90 percent probability bands, in percent. For each draw from the posterior, impulse responses are computed for the different regimes.

increase only moderately and not significantly, while uncertainty even increases on impact before the effect dies out. On the other hand, house prices, credit, exchange rates, and consumer confidence respond more strongly to a monetary policy shock compared with the acute phase. Nonetheless, the strong responses of these financial variables apparently do not markedly transmit to the real economy, i.e., output and prices, during the recovery phase.

The responses during normal recessions and expansions both show similar dynamics and are comparable to those of the normal-times regime presented in figure 1, but the responses of most variables turn out stronger in recessions than in expansions.[21] The effects in both normal-times regimes set in with larger delays and are often smaller than those in the acute phase, while being in parts stronger and more significant than those in the recovery phase of financial crises.

These differences in the role of monetary policy shocks across regimes are largely reflected in the forecast error variance decompositions, presented in table 3. Again, we observe that monetary policy shocks typically explain a larger share of forecast error variance for horizons of up to two years during financial crises compared with normal times, but differences across subperiods emerge. During the acute phase of financial crises, monetary policy shocks explain large fractions of forecast error in GDP, consumer prices, confidence, uncertainty, and particularly share prices. The shares are higher at a one-year horizon than in all other regimes, reflecting our result that the effects of monetary policy shocks set in rapidly during the acute phase. During the recovery phase, monetary policy shocks explain high shares of forecast error variation in house prices, exchange rates, confidence, and uncertainty, but less so in real variables. In normal recessions, monetary policy shocks explain considerably higher shares of forecast error variances compared with normal expansions for most variables, but somewhat lower shares compared with financial crises.

[21] Weise (1999), Garcia and Schaller (2002), Peersman and Smets (2002), and Lo and Piger (2005) also find that monetary policy is more effective during recessions than expansions, whereas more recent studies by Caggiano, Castelnuovo, and Groshenny (2014) and Tenreyro and Thwaites (2016) find the opposite.

Table 3. Forecast Error Variance Explained by the Monetary Policy Shock, Four Regimes

	Financial Crisis			
	Acute Phase		Recovery Phase	
	One Year	Two Years	One Year	Two Years
GDP	19.3 (11.6, 28.9)	23.9 (14.1, 35.8)	5.8 (2.1, 14.8)	11.0 (4.3, 25.6)
CPI	11.8 (4.6, 21.7)	15.7 (5.0, 29.6)	2.0 (0.3, 9.5)	3.3 (0.6, 14.6)
House Prices	1.2 (0.1, 5.6)	3.7 (0.6, 12.5)	24.9 (9.3, 41.4)	34.0 (13.7, 55.6)
Interest Rates	55.2 (44.6, 65.4)	52.6 (42.7, 62.2)	82.3 (67.7, 90.7)	71.0 (54.8, 83.3)
Credit	6.5 (1.6, 15.5)	7.9 (1.3, 20.0)	8.3 (1.3, 26.1)	15.5 (2.2, 37.9)
Exchange Rate	4.2 (1.4, 10.0)	5.4 (1.8, 12.9)	15.0 (3.9, 32.7)	15.9 (4.3, 32.3)
Confidence	14.9 (6.6, 25.0)	15.9 (7.4, 25.9)	22.9 (9.2, 40.4)	26.8 (12.4, 44.9)
Share Prices	38.7 (29.0, 48.5)	41.6 (31.5, 51.0)	10.6 (2.8, 24.3)	11.8 (4.3, 25.7)
Uncertainty	19.7 (12.2, 27.7)	20.2 (12.6, 27.7)	37.6 (23, 51.2)	37.5 (23.2, 51.1)
	Normal Times			
	Recession		**Expansion**	
	One Year	**Two Years**	**One Year**	**Two Years**
GDP	8.7 (2.4, 18.5)	25.2 (10.3, 42.0)	1.2 (0.4, 2.9)	6.4 (2.8, 11.6)
CPI	1.4 (0.1, 6.3)	2.2 (0.4, 8.2)	1.2 (0.2, 3.2)	1.2 (0.2, 4.0)
House Prices	10.2 (3.8, 19.6)	26.7 (13.5, 42.9)	2.9 (1.1, 5.5)	9.3 (4.8, 14.9)
Interest Rates	72.7 (59.9, 82.7)	60.4 (46.6, 73.2)	85.2 (80.4, 89.4)	74.0 (66.9, 80.3)
Credit	1.8 (0.4, 6.0)	9.9 (2.5, 24.1)	0.6 (0.2, 2.4)	3.5 (1.1, 8.3)
Exchange Rate	5.1 (1.3, 13.2)	13.6 (4.2, 28.1)	2.2 (0.6, 5.5)	2.9 (0.8, 7.4)
Confidence	11.5 (3.8, 23.3)	23.5 (10.9, 37.5)	4.0 (2.1, 7.0)	12.7 (7.6, 18.7)
Share Prices	16.9 (6.5, 29.8)	24.6 (10.2, 42.2)	2.9 (0.8, 7.0)	6.3 (2.0, 13.1)
Uncertainty	6.2 (1.6, 13.9)	8.0 (2.3, 18.2)	1.6 (0.5, 3.2)	1.8 (0.6, 3.6)

Notes: Median share of the forecast error variance explained by the monetary policy shock, in percent. The sixteenth and eighty-fourth credibility sets are shown in parentheses.

The series of identified monetary policy shocks from the extended model look very similar to those from the baseline model and are shown in the online appendix. They indicate that expansionary monetary surprises occurred at different stages of the global financial crisis when compared across countries. Again, the United States shows a large expansionary shock which occurred at the beginning of the acute phase of the global financial crisis, and also the mean value of U.S. monetary policy shocks was expansionary during this period. By contrast, in most euro-area member states, monetary policy surprises were on average contractionary during the acute phase, with large contractionary shocks occurring at its beginning of the global financial crisis. In those euro-area countries which exhibited double-dip recessions, the monetary surprises tend to be on average contractionary during the recessionary period in 2008 and 2009, but expansionary during the period 2011 to 2013. Hence, monetary policy shocks as identified in our model might have constituted additional adversary factors to euro-area economies during the financial crisis in 2008, while during the second recession these shocks have likely contributed to the recovery in the euro area.

4.3 Discussion of the Results

Overall, we find that monetary policy has larger and quicker effects on the economy during financial crises than during normal times, and that monetary policy shocks are more important for macroeconomic dynamics during financial crises, as indicated by forecast error variance decompositions. Further, there is some heterogeneity over the course of financial crises. During the acute phase of a crisis, the effects of monetary policy shocks are particularly large and quicker than during other times.

The results suggest that it is advisable to ease monetary policy quickly during financial crises. The counterfactual results show that indeed in the United States—where this strategy was pursued—the downturn in GDP was significantly mitigated, while this was not the case in the euro area, where interest rate cuts started one year later than in the United States. At the beginning of the financial crisis, euro-area monetary policy shocks were contractionary for many member economies according to our model and only became expansionary during the second recession in 2012/13 contributing to the final recovery. In contrast to the large effects during the acute phase of financial crises, we observe only weak effects of monetary policy on output and prices, such that the possibilities to stabilize macroeconomic variables appear limited during this phase. However, monetary policy shocks are not entirely without effect during this period, as financial market variables—like house prices, credit, and exchange rates—do react significantly. Macroeconomic uncertainty even significantly increases for a short period following an expansionary monetary policy shock. Hence, an early tightening of monetary policy would run the risk of financial markets destabilization.

Our finding that the effects of monetary policy are strong during the acute phase of financial crises, but less so in the recovery phase, is in line with findings of related empirical studies. Ciccarelli, Maddaloni, and Peydró (2013) analyze the effects of monetary policy in the euro area in the period from 2002 to 2011. By extending the end of their sample recursively from 2007 until 2011, they assess the evolution of the effects of a monetary policy shock when moving from a non-crisis period to the global financial crisis. In line with our results, they find that monetary policy became more effective at stimulating economic activity during the acute phase of the global financial crisis, mainly via the credit channel. Dahlhaus (2017) focuses on the United States and studies the effects of monetary policy conditional on financial stress, which shows that monetary policy is more effective in stimulating the economy during a high-stress regime. By contrast, Bech, Gambacorta, and Kharroubi (2014) focus on the recovery period of financial crises by studying the effects of the monetary policy stance during recession episodes on the strength of the subsequent recovery for a panel of twenty-four advanced economies. They find that monetary policy is not effective in stimulating GDP growth during the recovery phase of a financial crisis, whereas it is effective outside financial crises.

In a broader sense, the results of our study relate to numerous contributions in the literature that deal with the transmission of monetary policy under conditions typically prevailing during financial crises, such as credit constraints, high levels of uncertainty and financial stress, or the presence of large shocks. In the following, we discuss the potential mechanisms underlying our results in light of the findings in this literature.

Larger effects of monetary policy during financial crises and, in particular, during the acute phase are plausible if the central bank is able to ease the adverse effects of financial crises. In particular, firms and private households are more likely to be credit constrained during financial crises because of a decrease in the value of their financial assets and losses of collateral. In this situation, monetary policy may reduce the external finance premium by easing these constraints via the financial accelerator (Bernanke and Gertler 1995; Bernanke, Gertler, and Gilchrist 1999).

Further, our results suggest that the large responses of output during financial crises might relate to the strong effects monetary policy has on uncertainty, confidence, and share prices during this regime. Indeed, various contributions in the literature argue in favor of monetary policy transmission channels via these variables. Basu and Bundick (2017) show in a DSGE model with countercyclical markups and sticky prices that monetary policy can play a key role in offsetting the negative impact of uncertainty shocks. Bekaert, Hoerova, and Lo Duca (2013) provide empirical evidence for the effectiveness of expansionary monetary policy in decreasing risk aversion and uncertainty. By reducing uncertainty, monetary policy can stimulate output and inflation in various ways. First, lower uncertainty can lead to higher output because investment revives as investors receive new information (Bloom 2009). Second, a lower degree of uncertainty can stimulate credit supply and thus restore the credit channel of monetary policy. In particular, a lower degree of uncertainty should decrease liquidity risk and ease refinancing in interbank markets, which improves lending conditions for banks (Buch, Buchholz, and Tonzer 2014). Banks also have less need to retain capital for reasons of self-insurance and can thus extend their lending (Valencia 2017). Third, a reduction in uncertainty can improve consumer sentiment by enhancing the ability of agents to make probabilistic assessments about future events (Ilut and Schneider 2014). A monetary policy expansion can also raise consumer confidence directly by providing signals about future economic prospects (Bachmann and Sims 2012; Barsky and Sims 2012). In this context expansionary monetary policy might be interpreted as a sign that the central bank will prevent a further deepening of the crisis. An increase in consumer confidence can then restore the interest rate responsiveness of borrowing, investment, and spending on durables. Finally, an increase in share prices driven by a monetary policy expansion can increase the value of collateral, thereby contributing to a softening of credit constraints and to higher credit demand. An increase in share prices also makes it easier to finance investment by retaining profits.

Our finding that the co-movement between output and prices is much larger during financial crises than during normal times and in particular during the acute phase of financial crises is in line with the empirical literature on the Phillips curve. Several studies find a significant Phillips-curve relationship only when shocks are large. Roberts (2006) documents a flattening of the Phillips curve in the mid-1980s. This can potentially explain why it has become difficult since the mid-1980s to improve upon simple univariate inflation forecasts models, as documented by Atkeson and Ohanian (2001), Fisher, Liu, and Zhou (2002), and Orphanides and van Norden (2005). Stock and Watson (2007) systematically analyze reasons for the decline in the performance of activity-based inflation forecasts since the mid-1980s and argue that the decrease in the variance of activity measures during the Great Moderation is the reason for this. Further, Clark and McCracken (2006) argue that the sampling variability of inflation forecast comparison statistics is so large that evidence for a breakdown of the Phillips curve is not clear, despite the low accuracy of Phillips-curve-based inflation forecasts between the mid-1980s and the mid-2000s. All these studies use data prior to the 2008/09 Great Recession. However, the Great Recession might be a new episode of large macroeconomic shocks with output and inflation being highly volatile, so that a reemergence of a Phillips-curve relationship is plausible. Large monetary policy shocks can be one factor causing the increased co-movement between output and inflation according to our results, but other shocks—in particular, those that caused the Great Recession—are likely to be important as well.

Finally, the low effects on output and prices during the recovery phase of financial crises might be related to balance sheet adjustments and deleveraging by firms and financial institutions (Reinhart and Rogoff 2008). Monetary policy usually works via intertemporal substitution, but fewer people might be willing to increase their credit exposure in such periods. Hence, even highly expansionary monetary policy might not have a large effect on credit, output, and inflation.[22] Our result that share prices even decline and uncertainty even increases on impact following an expansionary monetary policy shock in this phase may mirror that such a shock may be regarded as confirmation of financial difficulties when the recovery is ongoing already for some time (Hubrich and Tetlow 2015). Continuous

[22] Agarwal et al. (2018) show that also monetary policy measures that were implemented during the Great Recession with the aim of stimulating household borrowing and spending by reducing banks' cost of funds have limited effects, because banks pass through credit expansion least to households that want to borrow the most.

monetary policy interventions could then be interpreted as signaling weak future fundamentals, thereby increasing uncertainty. This explanation would also be in line with the observation that exchange rate depreciation in response to an expansionary monetary policy shock is largest in the recovery phase of a financial crisis, but without leading to an increase in output.

Further, during the recovery phase of a financial crisis, factors that are important for the high effectiveness of monetary policy during the acute phase, such as the reduction of uncertainty and the stabilization of asset prices, may lose ground. At this late stage of a financial crisis, uncertainty has typically been mitigated and asset prices have stabilized substantially, with financial stress being much lower than during the acute phase. Thus, potential acceleration mechanisms in monetary policy transmission have already been exhausted.

4.4 Robustness Checks

We check the sensitivity of our results to a number of alternative specifications. In the following, we present the results for specifications which investigate the relevance of the zero lower bound and unconventional monetary policies. We also describe results of specifications, where we differentiate between euro-area and non-euro-area countries and where we use sign restrictions to identify monetary policy shocks. We show figures only for the baseline model with two regimes and for the three main variables output, prices, and the short-term interest rate. Figures for all variables and the extended model with four regimes are contained in the online appendix.

4.4.1 Zero Lower Bound and Quantitative Easing

The majority of the financial crises in our sample are associated with the global financial crisis that started in 2007: fifteen out of twenty financial crisis episodes took place around that period. In this crisis, several central banks reached the zero lower bound on nominal interest rates and therefore switched to unconventional monetary policy measures such as quantitative easing, for which we do not account in our baseline model.

We check whether our results are sensitive to accounting for unconventional monetary policy measure in two different ways. First, we use the shadow rate by Wu and Xia (2016) instead of the money market short-run interest rate for the United States, the United Kingdom, and the euro area as monetary policy instrument in our model. This measure is an estimate of the level of the short-term interest rate that would prevail in the absence of the zero lower bound. It is estimated based on forward rates and a Nelson-Siegel-Svensson yield-curve model, so it also reflects quantitative easing and other unconventional monetary policy measures.[23] The results of this specification are shown in figure 5, as dotted lines. The responses of both output and prices remain stronger during financial crises than during normal times, although the responses of both variables during financial crises are slightly weaker when using shadow interest rate estimates compared with the baseline. A similar pattern holds for the other six variables in the model. In the estimation over four regimes, results when using the shadow interest rate as monetary policy instrument are very similar to the baseline, as shown in the online appendix.

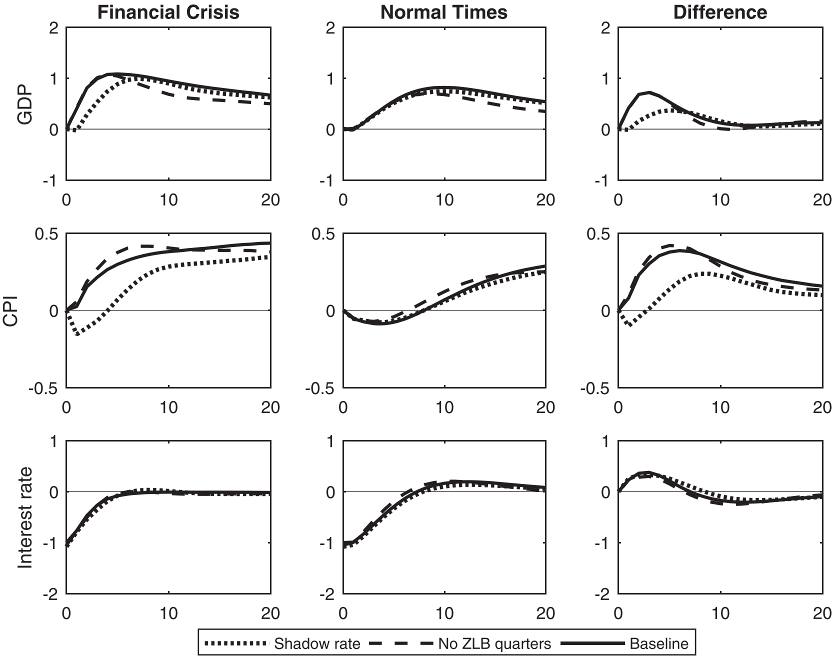
Second, we perform a more agnostic robustness check with respect to the zero lower bound and unconventional monetary policy measures. We identify all observations in which monetary policy reaches the zero lower bound, i.e., when the interest rate of the main monetary policy instrument is smaller than or equal to 0.3 percent. This definition yields 298 observations where monetary policy reaches the zero lower bound, among which 33 occur in financial crises. When we exclude these observations from our sample, our results are basically unchanged, as shown by the dashed lines in figure 5.

4.4.2 Euro-Area and Non-euro-area Samples

Eleven out of the twenty countries in our sample are members of the euro area. In our data set, the aggregate euro-area short-term

[23]Shadow interest rates are not available for other economies; therefore, we do not control for unconventional monetary policy measures in Japan, which is the only remaining economy in our sample where such measures have been applied for a longer period. However, our results do not change when we also exclude Japan from our sample.

Figure 5. Accounting for the Zero Lower Bound

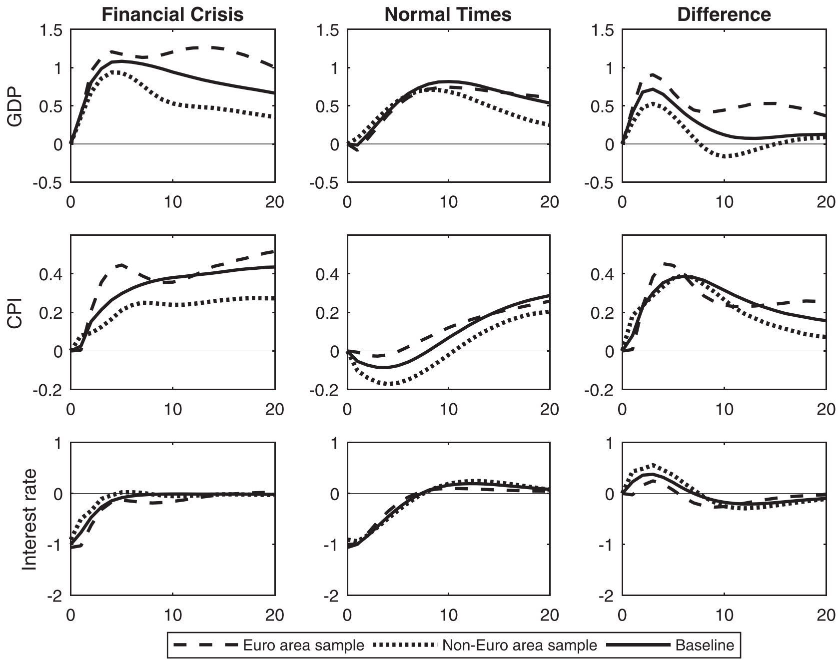


Notes: Solid lines show baseline estimation; dotted lines show estimation with the shadow interest rate as the monetary policy variable; dashed lines show results for the estimation where the 298 observations in which the zero lower bound was reached were excluded. See also notes of figure 1.

interest rate is thus used for all euro-area members from 1999 onwards (from 2001 in the case of Greece). While we opt for including these countries separately in order to exploit the cross-country dimension of the data and to be able to analyze differences across euro-area member states in section 4.1.2, the information that can be drawn from including individual euro-area countries in our panel might be limited given that they share a joint monetary policy.[24]

[24]The existing empirical literature analyzing monetary policy transmission within a panel VAR context typically follows a similar approach and includes euro-area countries as individual economies in the panel (Assenmacher-Wesche and Gerlach 2008; Goodhart and Hofmann 2008; Carstensen, Hülsewig, and Wollmerhäusen 2009; Calza, Monacelli, and Stracca 2013). Only Gambacorta, Hofmann, and Peersman (2014) take a different approach and consider the euro

Figure 6. Euro-Area and Non-euro-area Samples



Notes: Solid lines show baseline estimation; dashed (dotted) lines show estimation for the euro area (non-euro-area) subsample. See also notes of figure 1.

To assess whether our results are dominated by euro-area dynamics, we run two separate PVAR regressions for euro-area countries and for countries outside the euro area, respectively. We only have about half of the total observations for each subgroup, so estimation is only feasible for the baseline case with two regimes. Figure 6 shows the median impulse responses and probability bands for the euro-area sample (dashed lines) and the non-euro-area sample (dotted lines), together with the baseline estimates for the full sample, depicted as solid lines. The results are comparable to the baseline estimation and mostly similar for both subgroups. Output and prices

area as aggregate, at the cost of having only eight countries in their panel VAR, which for our purposes of pooling financial crisis episodes and distinguishing between different regimes would not provide enough data points.

react stronger and faster during financial crises in both the euro-area countries and the non-euro area countries. Differences between euro-area and non-euro-area countries are small for all other variables.

We also check whether our results are affected by the sovereign debt crisis that become severe in 2011 in several euro-area countries, leading to a double-dip recession in these economies. To exclude the sovereign debt crisis, we restrict our estimation period to the subsample from 1984 to 2010 and our results remain strongly robust, as shown in the online appendix.

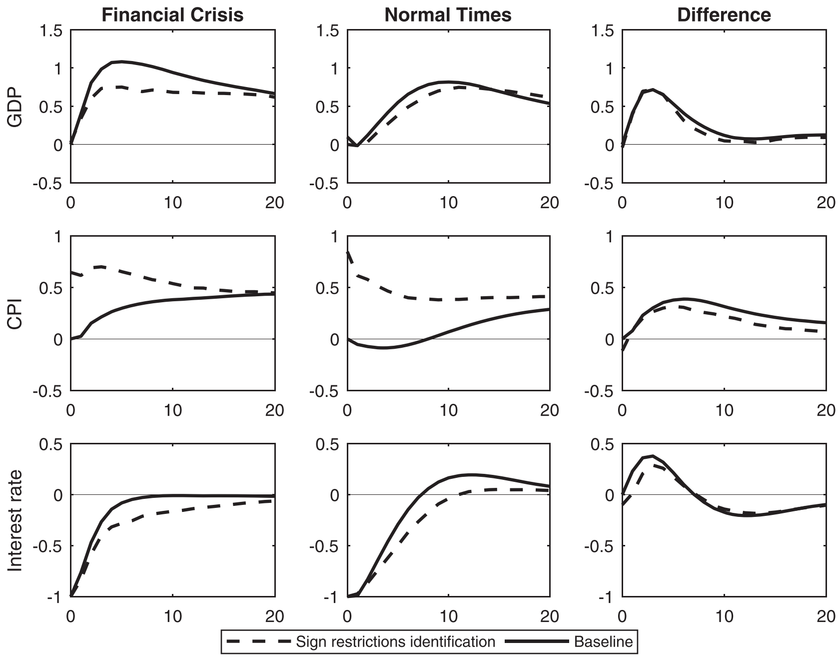
4.4.3 Sign Restrictions

Finally, we check whether our results are robust when we use sign restrictions instead of the recursive identification to identify monetary policy shocks. The sign-restriction approach has the advantage that it allows for non-zero contemporaneous responses by all variables in the VAR. We remain agnostic about the response of output, and we restrict the impulse response of the short-term interest rate to being negative during the first four quarters and the response of prices to being positive between the second and fourth quarter. We follow Sá, Towbin, and Wieladek (2014) and do not restrict the CPI response in the first quarter to allow for a sluggish response by this variable. Figure 7 shows that the results are similar to the results obtained based on the recursive identification. Output and prices react faster and stronger during financial crises. The results for the other six variables are similar to our baseline results as well: we find a stronger increase in confidence and share prices as well as a larger reduction in uncertainty during financial crises. Also in the estimation over four regimes, results are very similar when using sign-restriction identification, as shown in the online appendix.

5. Conclusion

In this paper, we analyze the effects of monetary policy shocks during financial crises. We find that monetary policy shocks have significantly larger effects on output and prices during financial crises than during normal times and that these effects occur more instantaneously. We also observe significantly larger effects on several other variables, including credit, asset prices, uncertainty, and consumer

Figure 7. Identification via Sign Restrictions



Notes: Solid lines show baseline estimation; dashed lines show results for the estimation, where the monetary policy shock is identified via sign restrictions.

confidence, suggesting that these variables play an important role in the monetary policy transmission during financial crises. However, we find that there is time variation of the effects of monetary policy shocks during financial crises. While monetary policy shocks have large effects on output and prices during the acute phase of a financial crisis when the economy is in recession, the effects are significantly weaker during the subsequent recovery phase.

Further, we find some heterogeneity of monetary policy actions during the global financial crisis of 2008/09 across countries. In economies like the United States, in which interest rates were lowered directly at the beginning of the crisis, monetary policy was able to mitigate the biggest drop in output to some extent, but less so in other economies like the euro area, where the central bank reacted later. Our results suggest that monetary policy is highly effective during financial crises, probably because it can reduce some of the adverse characteristics of financial crises like high uncertainty, low confidence, and low asset prices. We discuss several papers that support this view, but leave a more detailed analysis of the relevance of different transmission channels of monetary policy during financial crises for future research.

References

Agarwal, S., S. Chomsisengphet, N. Mahoney, and J. Stroebel. 2018. "Do Banks Pass through Credit Expansions to Consumers Who Want to Borrow? *Quarterly Journal of Economics* 133 (1): 129–90.

Altunbasa, Y., L. Gambacorta, and D. Marques-Ibanez. 2014. "Does Monetary Policy Affect Bank Risk?" *International Journal of Central Banking* 10 (1, March): 95–135.

Assenmacher-Wesche, K., and S. Gerlach. 2008. "Monetary Policy, Asset Prices and Macroeconomic Conditions: A Panel-VAR Study." Working Paper No. 149, National Bank of Belgium.

Atkeson, A., and L. E. Ohanian. 2001. "Are Phillips Curves Useful for Forecasting Inflation?" *Quarterly Review* (Federal Reserve Bank of Minneapolis) 25 (1): 2–11.

Bachmann, R., and E. Sims. 2012. "Confidence and the Transmission of Government Spending Shocks." *Journal of Monetary Economics* 59 (3): 235–49.

Barigozzi, M., A. M. Conti, and M. Luciani. 2014. "Do Euro Area Countries Respond Asymmetrically to the Common Monetary Policy?" *Oxford Bulletin of Economics and Statistics* 76 (5): 693–714.

Barsky, R. B., and E. R. Sims. 2012. "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence." *American Economic Review* 102 (4): 1343–77.

Basu, S., and B. Bundick. 2017. "Uncertainty Shocks in a Model of Effective Demand." *Econometrica* 85 (3): 937–58.

Bech, M., L. Gambacorta, and E. Kharroubi. 2014. "Monetary Policy in a Downturn: Are Financial Crises Special?" *International Finance* 17 (1): 99–119.

Bekaert, G., M. Hoerova, and M. Lo Duca. 2013. "Risk, Uncertainty and Monetary Policy." *Journal of Monetary Economics* 60 (7): 771–88.

Bernanke, B. S. 1983. "Irreversibility, Uncertainty, and Cyclical Adjustment." *Quarterly Journal of Economics* 98 (1): 85–106.

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

Bernanke, B. S., and M. Gertler. 1995. "Inside the Black Box: The Credit Channel of Monetary Policy Transmission." *Journal of Economic Perspectives* 9 (4): 27–48.

Bernanke, B. S., M. Gertler, and S. Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework." In *Handbook of Macroeconomics*, Vol. 1C, ed. J. B. Taylor and M. Woodford, 1341–93 (chapter 21). Elsevier.

Bloom, N. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77 (3): 623–85.

Bloom, N., S. Bond, and J. V. Reenen. 2007. "Uncertainty and Investment Dynamics." *Review of Economic Studies* 74 (2): 391–415.

Borio, C., and B. Hofmann. 2017. "Is Monetary Policy Less Effective when Interest Rates Are Persistently Low?" BIS Working Paper No. 628.

Bouis, R., L. Rawdanowicz, J.-P. Renne, S. Watanabe, and A. K. Christensen. 2013. "The Effectiveness of Monetary Policy since the Onset of the Financial Crisis." OECD Economics Department Working Paper No. 1081.

Buch, C. M., M. Buchholz, and L. Tonzer. 2014. "Uncertainty and International Banking." Mimeo.

Caggiano, G., E. Castelnuovo, and N. Groshenny. 2014. "Uncertainty Shocks and Unemployment Dynamics in U.S. Recessions." *Journal of Monetary Economics* 67 (October): 78–92.

Calza, A., T. Monacelli, and L. Stracca. 2013. "Housing Finance and Monetary Policy." *Journal of the European Economic Association* 11 (s1): 101–22.

Canova, F., and G. De Nicoló. 2002. "Monetary Disturbances Matter for Business Fluctuations in the G-7." *Journal of Monetary Economics* 49 (6): 1131–59.

Carstensen, K., O. Hülsewig, and T. Wollmershäuser. 2009. "Monetary Policy Transmission and House Prices: European Cross-Country Evidence." CESifo Working Paper No. 2750.

Cesa-Bianchi, A., M. H. Pesaran, and A. Rebucci. 2018. "Uncertainty and Economic Activity: A Multi-country Perspective." NBER Working Paper No. 24325.

Chen, Q., A. Filardo, D. He, and F. Zhu. 2016. "Financial Crisis, US Unconventional Monetary Policy and International Spillovers." *Journal of International Money and Finance* 67 (October): 62–81.

Christiano, L. J., M. Eichenbaum, and C. L. Evans. 1999. "Monetary Policy Shocks: What Have We Learned and to What End?" In *Handbook of Macroeconomics*, Vol. 1A, 1st Edition, ed. J. B. Taylor and M. Woodford, 65–148 (chapter 2). Elsevier.

Ciccarelli, M., A. Maddaloni, and J.-L. Peydró. 2013. "Heterogeneous Transmission Mechanism: Monetary Policy and Financial Fragility in the Eurozone." *Economic Policy* 28 (75): 459–512.

Clark, T. E., and M. W. McCracken. 2006. "The Predictive Content of the Output Gap for Inflation: Resolving In-Sample and Out-of-Sample Evidence." *Journal of Money, Credit and Banking* 38 (5): 1127–48.

Cogley, T., and T. J. Sargent. 2005. "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US." *Review of Economic Dynamics* 8 (2): 262–302.

Coibion, O. 2012. "Are the Effects of Monetary Policy Shocks Big or Small?" *American Economic Journal: Macroeconomics* 4 (2): 1–32.

Corsetti, G., J. B. Duarte, and S. Mann. 2018. "One Money, Many Markets—A Factor Model Approach to Monetary Policy in the Euro Area with High-Frequency Identification." CFM Discussion Paper No. 1805.

Dahlhaus, T. 2017. "Conventional Monetary Policy Transmission during Financial Crises: An Empirical Analysis." *Journal of Applied Econometrics* 32 (2): 401–21.

Del Negro, M., and F. Schorfheide. 2011. "Bayesian Macroeconometrics." In *The Oxford Handbook of Bayesian Econometrics*, ed. J. Geweke, G. Koop, and H. V. Dijk, 293–389. Oxford University Press.

Dixit, A., and R. Pindyck. 1994. *Investment under Uncertainty.* Princeton, NJ: Princeton University Press.

Fisher, J. D., C. T. Liu, and R. Zhou. 2002. "When Can We Forecast Inflation?" *Economic Perspectives* (Federal Reserve Bank of Chicago) 26 (1): 30–42.

Fry-McKibbin, R., and J. Zheng. 2016. "Effects of US Monetary Policy Shocks during Financial Crises — A Threshold Vector Autoregression Approach." *Applied Economics* 48 (59): 5802–23.

Gambacorta, L., B. Hofmann, and G. Peersman. 2014. "The Effectiveness of Unconventional Monetary Policy at the Zero Lower Bound: A Cross-Country Analysis." *Journal of Money, Credit and Banking* 46 (4): 615–42.

Garcia, R., and H. Schaller. 2002. "Are the Effects of Monetary Policy Asymmetric?" *Economic Inquiry* 40 (1): 102–19.

Georgiadis, G. 2014. "Towards an Explanation of Cross-Country Asymmetries in Monetary Transmission." *Journal of Macroeconomics* 39 (Part A): 66–84.

Giannone, D., M. Lenza, H. Pill, and L. Reichlin. 2012. "The ECB and the Interbank Market." *Economic Journal* 122 (564): F467–F486.

Goodhart, C., and B. Hofmann. 2008. "House Prices, Money, Credit, and the Macroeconomy." *Oxford Review of Economic Policy* 24 (1): 180–205.

Harding, D., and A. Pagan. 2002. "Dissecting the Cycle: A Methodological Investigation." *Journal of Monetary Economics* 49 (2): 365–81.

Hubrich, K., and R. J. Tetlow. 2015. "Financial Stress and Economic Dynamics: The Transmission of Crises." *Journal of Monetary Economics* 70 (March): 100–115.

Ilut, C., and M. Schneider. 2014. "Ambiguous Business Cycles." *American Economic Review* 104 (8): 2368–99.

Jiménez, G., S. Ongena, J.-L. Peydró, and J. Saurina. 2014. "Hazardous Times for Monetary Policy: What Do Twenty-Three Million Bank Loans Say about the Effects of Monetary Policy on Credit Risk-Taking?" *Econometrica* 82 (2): 463–505.

Koop, G., and D. Korobilis. 2010. *Bayesian Multivariate Time Series Methods for Empirical Macroeconomics.* Now Publishers Inc.

Laeven, L., and F. Valencia. 2013. "Systemic Banking Crises Database." *IMF Economic Review* 61 (2): 225–70.

Lo, M. C., and J. Piger. 2005. "Is the Response of Output to Monetary Policy Asymmetric? Evidence from a Regime-Switching Coefficients Model." *Journal of Money, Credit and Banking* 37 (5): 865–86.

Lütkepohl, H. 2013. "Vector Autoregressive Models." In *Handbook of Research Methods and Applications in Empirical Macroeconomics*, ed. N. Hashimzade and M. A. Thornton, 139–64 (part II, chapter 6). Edward Elgar Publishing.

Mishkin, F. S. 2009. "Is Monetary Policy Effective during Financial Crises?" *American Economic Review* 99 (2): 573–77.

Morgan, D. P. 1993. "Asymmetric Effects of Monetary Policy." *Economic Review* (Federal Reserve Bank of Kansas City) 78 (Second Quarter): 21–33.

Nickell, S. J. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49 (6): 1417–26.

Orphanides, A., and S. van Norden. 2005. "The Reliability of Inflation Forecast Based on Output Gap Estimates in Real Time." *Journal of Money, Credit and Banking* 37 (3): 583–600.

Pain, N., C. Lewis, T.-T. Dang, Y. Jin, and P. Richardson. 2014. "OECD Forecasts during and after the Financial Crisis: A Post Mortem." OECD Economics Department Working Paper No. 1107.

Peersman, G., and F. Smets. 2002. "Are the Effects of Monetary Policy in the Euro Area Greater in Recessions than in Booms?" In *Monetary Transmission in Diverse Economies*, ed. L. Mahadeva and P. Sinclair, 28–48 (chapter 2). Cambridge University Press.

Pesaran, M. H., and R. Smith. 1995. "Estimating Long-run Relationships from Dynamic Heterogeneous Panels." *Journal of Econometrics* 68 (1): 79–113.

Rajan, R. G. 2005. "Has Financial Development Made the World Riskier?" In *The Greenspan Era: Lessons for the Future.* Proceedings of the 2005 Economic Policy Symposium sponsored by the Federal Reserve Bank of Kansas City, held in Jackson Hole, Wyoming, August 25–27.

Reinhart, C. M., and K. S. Rogoff. 2008. "Is the 2007 US Subprime Financial Crisis So Different? An International Historical Comparison." *American Economic Review* 98 (2): 339–44.

Roberts, J. M. 2006. "Monetary Policy and Inflation Dynamics." *International Journal of Central Banking* 2 (3, September): 193–230.

Sá, F., P. Towbin, and T. Wieladek. 2014. "Capital Inflows, Financial Structure and Housing Booms." *Journal of the European Economic Association* 12 (2): 522–46.

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.

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

Towbin, P., and S. Weber. 2013. "Limits of Floating Exchange Rates: The Role of Foreign Currency Debt and Import Structure." *Journal of Development Economics* 101 (March): 179–94.

Uhlig, H. 1994. "What Macroeconomists Should Know about Unit Roots: A Bayesian Perspective." *Econometric Theory* 10 (3–4): 645–71.

———. 2005. "What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure." *Journal of Monetary Economics* 52 (2): 381–419.

Valencia, F. 2017. "Aggregate Uncertainty and the Supply of Credit." *Journal of Banking and Finance* 81 (August): 150–65.

Vavra, J. 2014. "Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation." *Quarterly Journal of Economics* 129 (1): 215–58.

Weise, C. L. 1999. "The Asymmetric Effects of Monetary Policy: A Nonlinear Vector Autoregression Approach." *Journal of Money, Credit and Banking* 31 (1): 85–108.

Williams, J. C. 2014. "Monetary Policy at the Zero Lower Bound: Putting Theory into Practice." Working Paper No. 2 (January), Hutchins Center, Brookings Institution.

Wu, J. C., and F. D. Xia. 2016. "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound." *Journal of Money, Credit and Banking* 48 (2–3): 253–91.

Online Appendix to Monetary Policy during Financial Crises: Is the Transmission Mechanism Impaired?

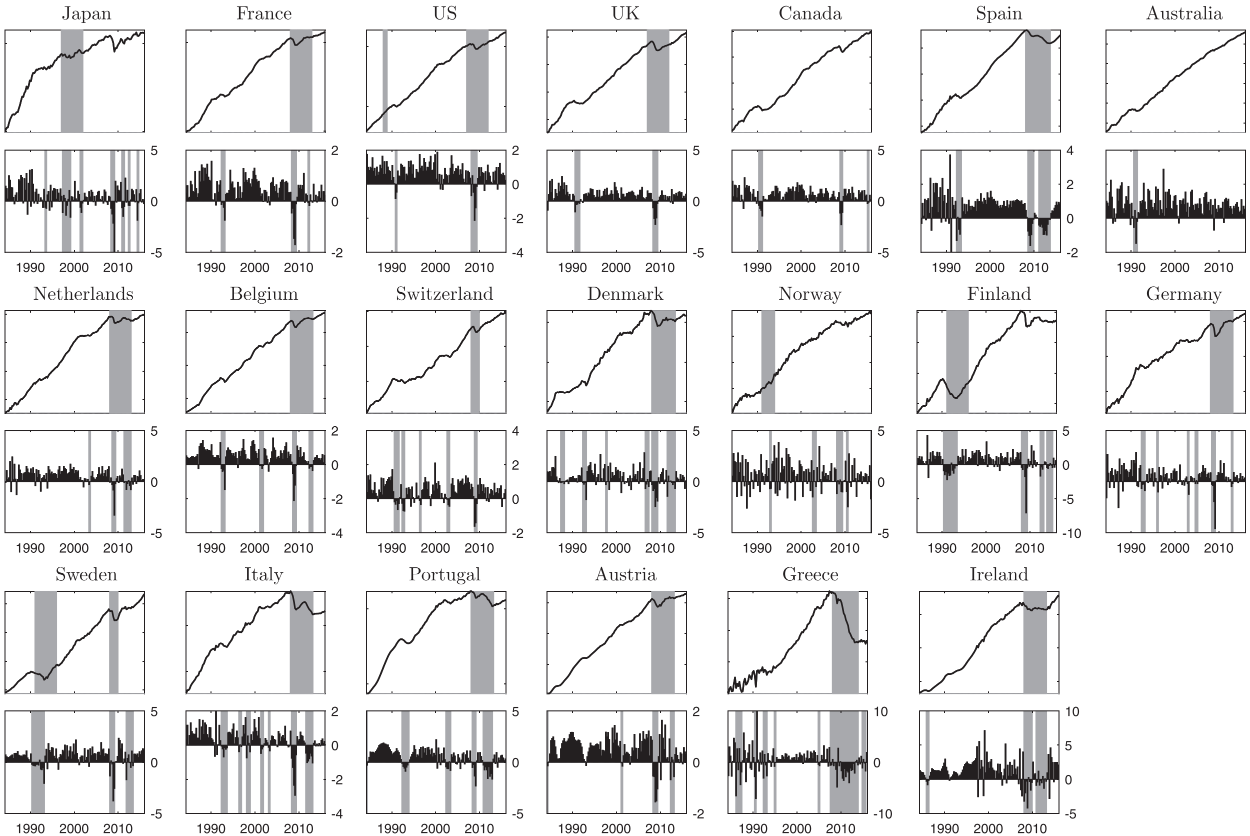
Nils Jannsen,[a] Galina Potjagailo,[a,b] and Maik H. Wolters[a,c,d]
[a]Kiel Institute for the World Economy
[b]University of Kiel
[c]University of Jena
[d]IMFS at Goethe-University Frankfurt

In this complementary appendix we provide additional material, which gives a more detailed picture of our results, but which is of secondary importance to the main results presented in the paper. In table A1, we show information on the data sources for our main endogenous variables and on the twenty countries included in the panel analysis. In figure A1 we present the financial crisis and recession episodes for each country. Figures A2 and A3 present additional results for the extended interacted panel VAR model with four regimes. Finally, figures A4–A8 show detailed results for the robustness checks that we discuss in the paper.

Table A1. Data Sources and Description

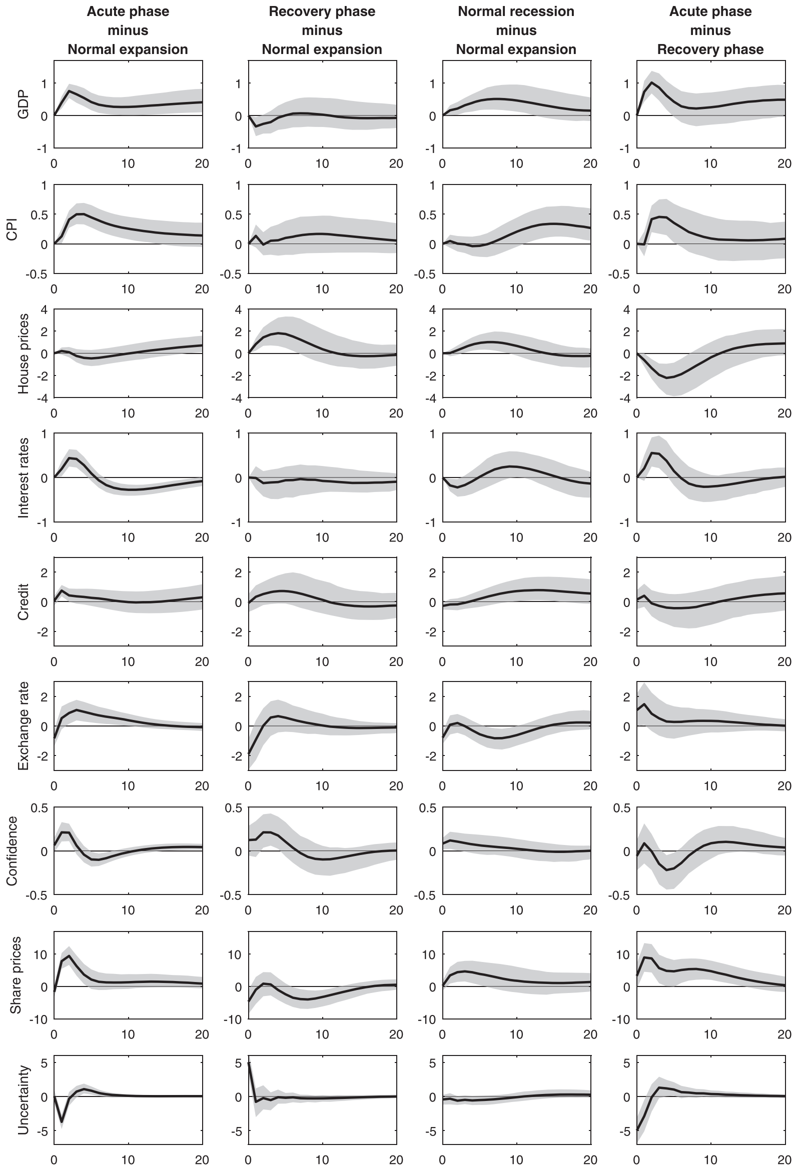
Variable	Source	Remarks
Gross Domestic Product	OECD Economic Outlook	Seasonally adjusted (SA). At 2005 market prices.
CPI, All Items	OECD Main Economic Indicators	SA (seasonal filter). Index including all items, 2010 = 100.
Short-Term Interest Rate	OECD Economic Outlook	Three-month short-term interest rate, Greece from 1995, Ireland from 1990.
House Prices	Federal Reserve Bank of Dallas	SA (seasonal filter). Portugal from 1988, Austria from 2000, Greece from 1997.
Total Credit to Private Sector	Bank for International Settlements	SA (seasonal filter). Total amount of nominal credit (i.e., loans and debt securities) provided by domestic banks to non-financial corporations, households, and non-profit institutions serving households. In national currency. See http://www.bis.org/statistics/credtopriv.htm).
Effective Exchange Rate	OECD Economic Outlook	SA (seasonal filter). Greece from 1995, Ireland from 1990.
Consumer Confidence Indicator	National Sources	SA (seasonal filter). Standardized. From 1985 for most countries. Norway from 1992, Sweden from 1990, Austria from 1996.
Share Prices	OECD Main Economic Indicators	SA (seasonal filter). Spain, Belgium, Greece from 1985, Norway from 1986, Portugal from 1988.
Stock Volatility	Based on Daily Major Stock Market Indexes from Datastream	SA (seasonal filter), own calculation as in Cesa-Bianchi, Pesaran, and Rebucci (2018), Spain, Finland, Greece from 1988, Portugal from 1990.
Countries Included	Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, United States.	

Figure A1. Financial Crises and Recessions Data for Twenty OECD Economies



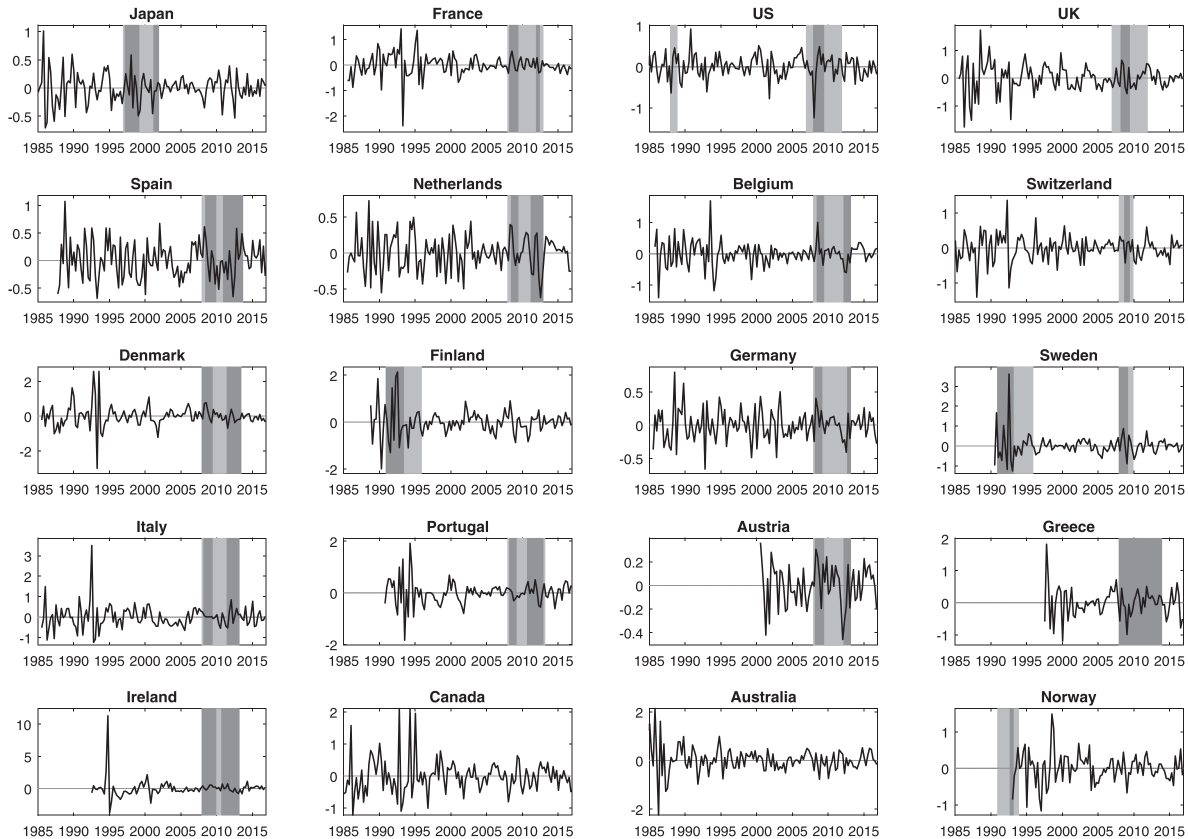
Notes: The figures in the upper part show log real GDP and the figures in the lower part show quarterly (non-annualized) real GDP growth. Shaded areas represent identified financial crisis episodes in the upper part and identified recessions in the lower part. The outlier in GDP growth in Ireland in 2015:Q1 was replaced by the median of the five previous observations of the time series.

Figure A2. Differences between Acute and Recovery Phases of Financial Crisis and Normal-Times Regimes



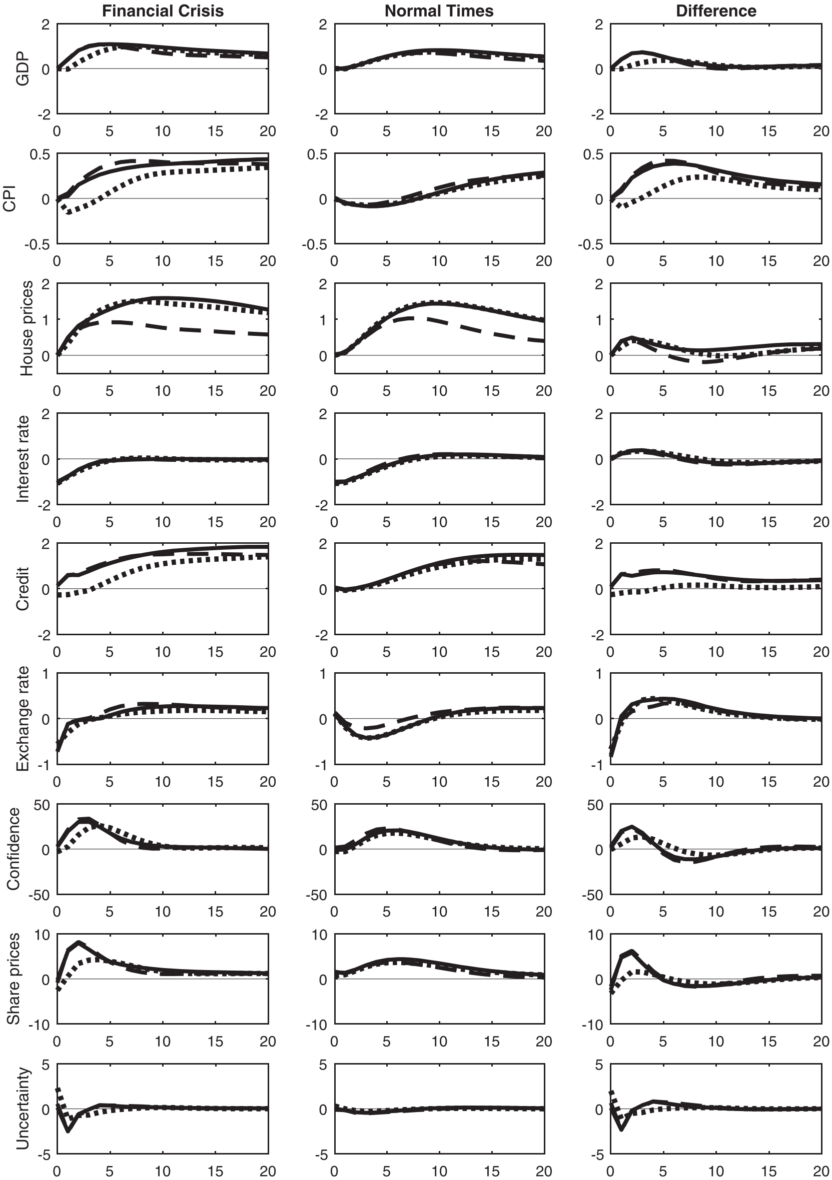
Notes: Columns 1–3 show median differences between the impulse responses evaluated in the acute phase, the recovery phase, and normal recessions, respectively, relative to the impulse responses evaluated in the normal expansion regime, in percentage points. Column 4 shows the difference between the impulse responses evaluated in the acute phase of a financial crisis and in normal recessions, in percentage points. Gray areas represent 90 percent probability bands. These results were obtained by calculating, for each draw, the difference between the impulse responses evaluated at two respective regimes and by taking the median and percentile values of these differences.

Figure A3. Estimated Monetary Policy Shocks



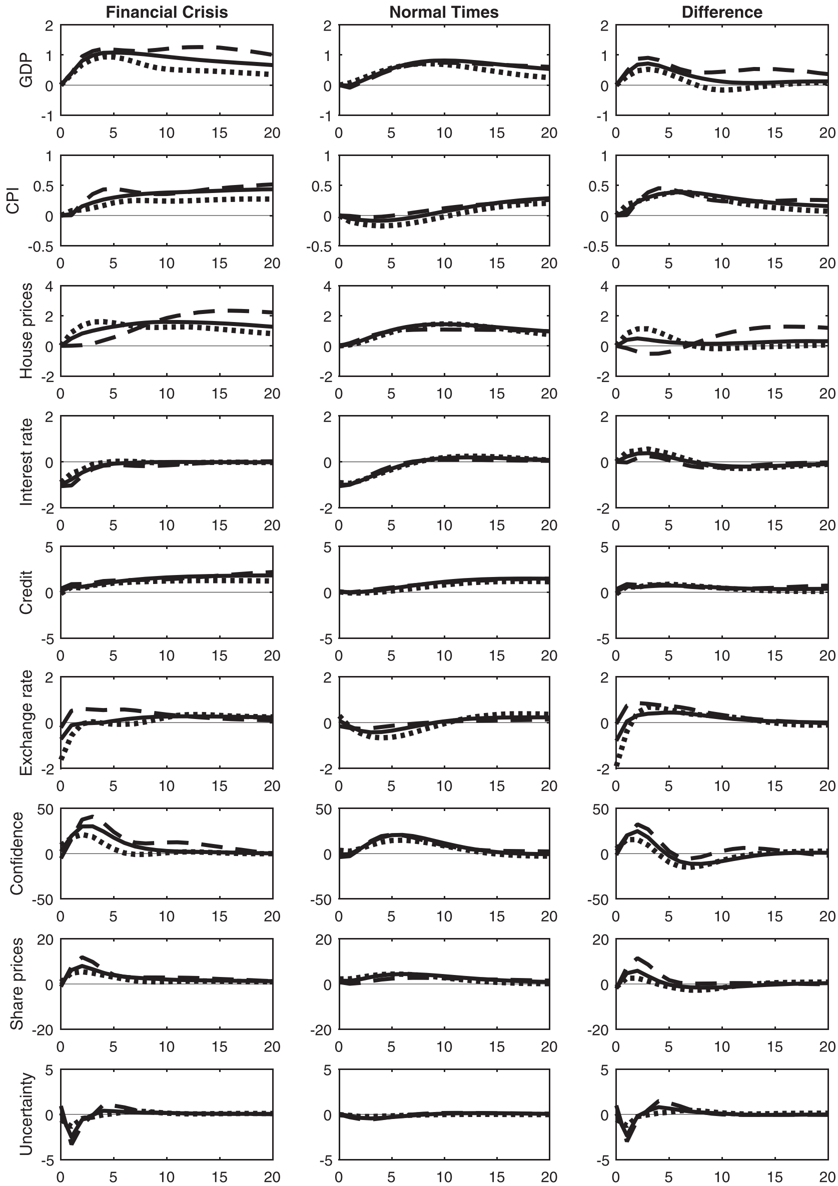
Notes: Solid lines show the residuals corresponding to the interest rate equation of each country from the structural form of the IPVAR of equation (2), extended model with four regimes. Light gray bars show the financial crisis episodes for each country; dark gray bars show recession episodes that occur during financial crises (acute phase).

Figure A4. Accounting for the Zero Lower Bound: Full Results



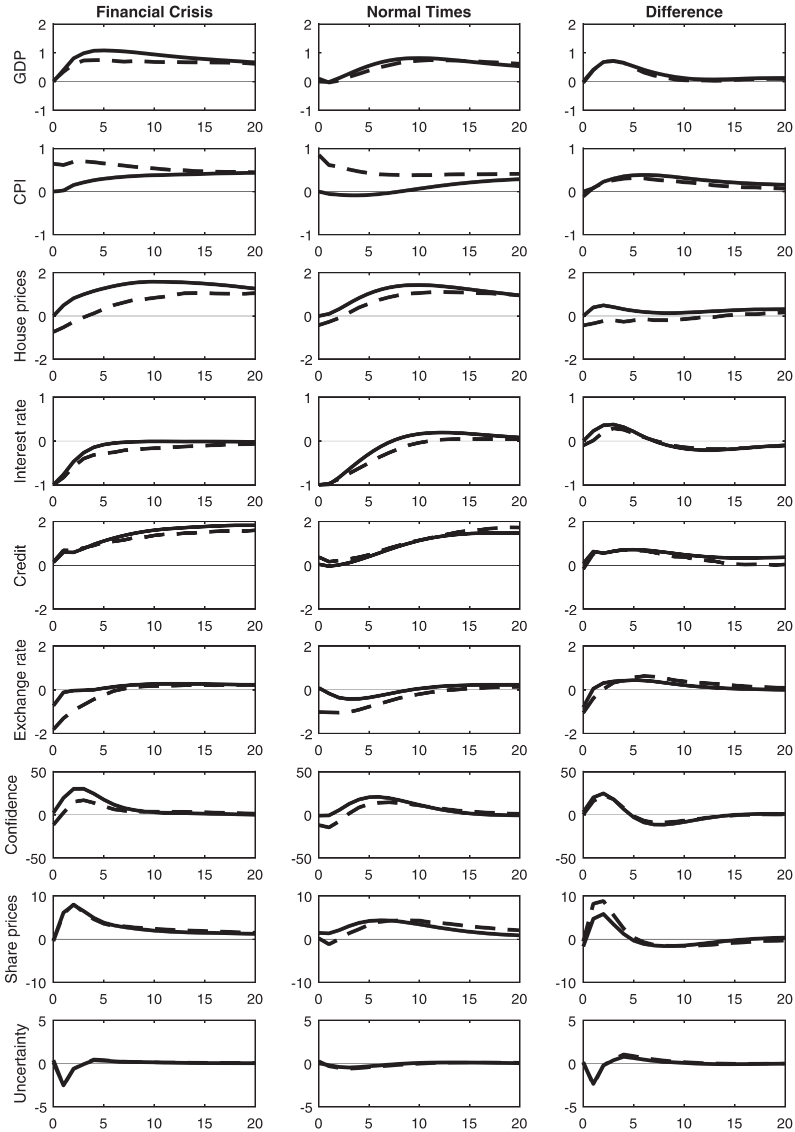
Notes: Solid lines show results for the baseline estimation; dotted lines show results for the estimation with the shadow interest rate as monetary policy variable; dashed lines show results for the estimation where the 298 observations in which the zero lower bound was reached were excluded. Median impulse responses to a 100 basis point expansionary monetary policy shock during financial crises and normal times are in percent; median differences between regimes are in percentage points.

Figure A5. Euro-Area and Non-euro-area Samples: Full Results



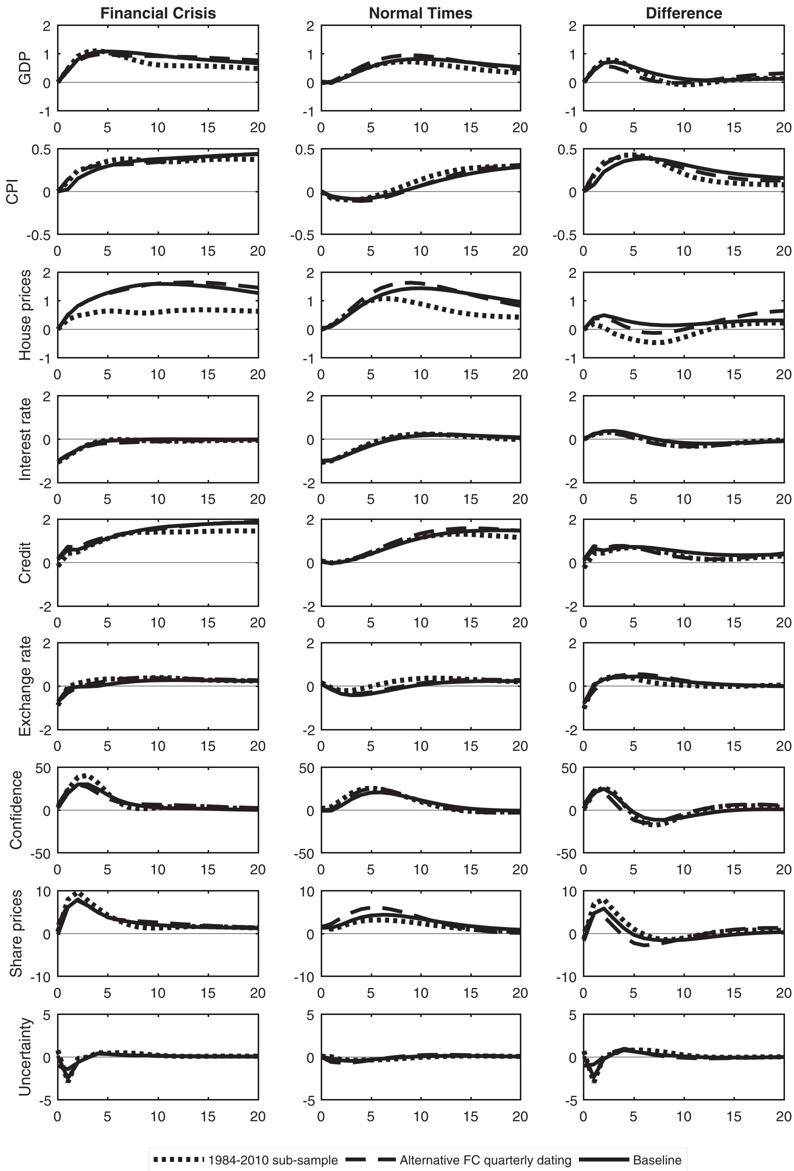
Notes: Solid lines show results from the baseline estimation; dashed (dotted) lines show results for the estimation over the euro-area (non-euro-area) subsample. Median impulse responses to a 100 basis point expansionary monetary policy shock during financial crises and normal times are in percent; median differences between regimes are in percentage points.

Figure A6. Identification via Sign Restrictions: Full Results



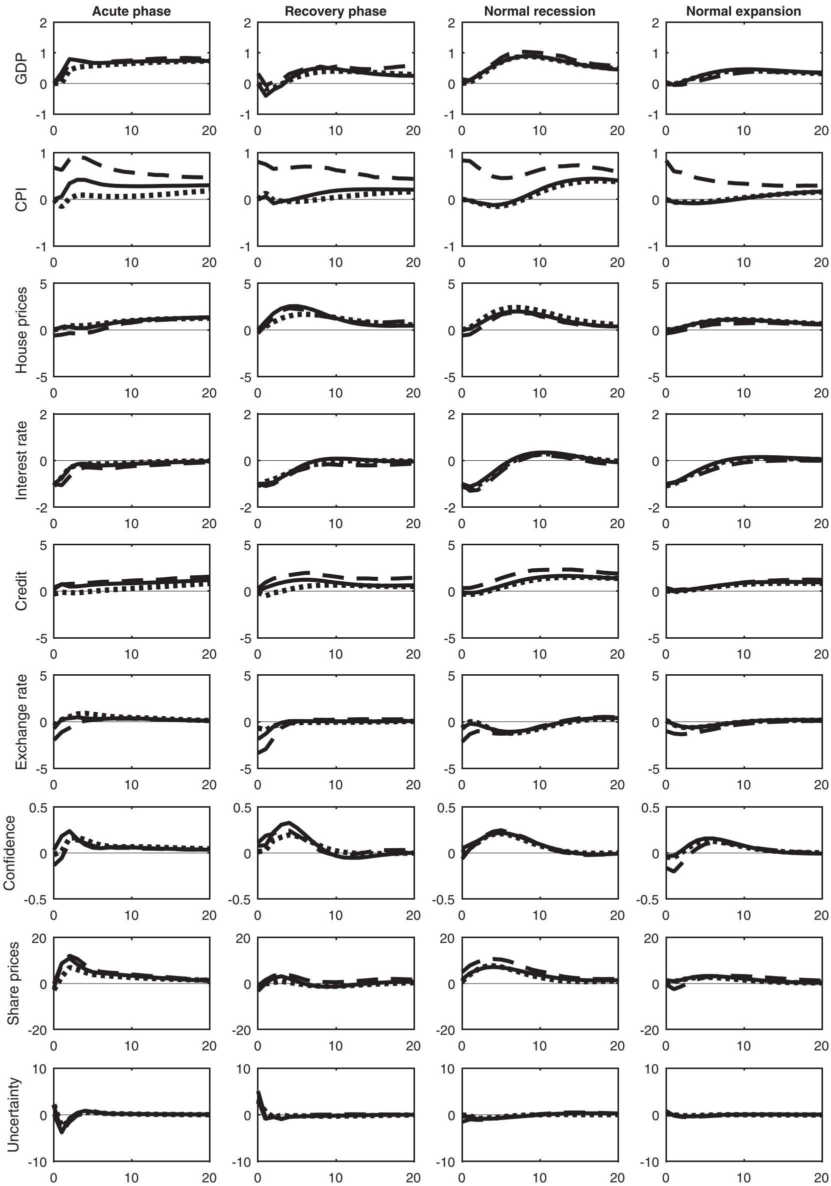
Notes: Solid lines show results from the baseline estimation; dashed lines show results for the estimation, where the monetary policy shock is identified via sign restrictions. Median impulse responses to a 100 basis point expansionary monetary policy shock during financial crises and normal times are in percent; median differences between regimes are in percentage points. In the robustness check, the recursively estimated coefficients are evaluated at different values of the interaction terms and reduced-form parameters are calculated for each draw from the posterior. Then, for each parameter draw and regime, rotations of the impulse responses are generated until 100 rotations satisfy the sign restrictions. The median over the accepted rotations is then retained for each parameter draw and another median is calculated over all parameter draws.

Figure A7. Additional Robustness Checks



Notes: Solid lines show results from the baseline estimation; dashed lines show results for the estimation with an alternative quarterly dating of financial crisis episodes (beginning in the last quarter of the first financial crisis year and ending in the first quarter of the last year). Dotted lines show results for the estimation over the 1984–2010 subsample which excludes the double-dip recession period after the global financial crisis. Median impulse responses to a 100 basis point expansionary monetary policy shock during financial crises and normal times are in percent; median differences between regimes are in percentage points.

Figure A8. Four Regimes: Estimation with Shadow Interest Rates and Identification via Sign Restrictions



Notes: Solid lines show the baseline estimation over four regimes; dashed lines show results for the estimation, where the monetary policy shock is identified via sign restrictions; dotted lines show results for the estimation with the shadow interest rate as monetary policy variable (with recursive identification as in the baseline). Median impulse responses to a 100 basis point expansionary monetary policy shock are in percent. For each draw from the posterior, impulse responses are computed for the different regimes.

Spillovers from the ECB's Non-standard Monetary Policy Measures on Southeastern Europe*

Isabella Moder
European Central Bank

This paper is the first to comprehensively assess the impact of the euro area's non-standard monetary policy measures on southeastern Europe. The outcomes of bilateral BVAR models suggest that the ECB's non-standard monetary policy measures have had pronounced price effects on all, and output effects on approximately half, of the countries in southeastern Europe. While I find evidence that exports have posed as relevant transmission channels in most cases, the role of the interbank market rate as a channel of shock transmission is less clear. Furthermore, the results suggest that exchange rates' responses have been relatively muted.

JEL Codes: C11, C32, E52, F42.

1. Introduction

Since October 2008 the European Central Bank (ECB) has introduced a number of non-standard monetary policy measures, which are unprecedented in nature, scope, and magnitude, and have ranged from significant changes in the operational framework to large bond purchasing programs. Assessing potential spillovers from monetary policy measures of advanced economies has become important in a globalized world, and this incorporates not only potential

*I would like to thank the editor, two anonymous referees, Martin Feldkircher, Philipp Harms, Éva Katalin Polgár, Georgios Georgiadis, the members of the ESCB's International Relations Committee, and participants of the JGU Mainz Brown Bag Seminar and of an internal ECB seminar for helpful comments and suggestions. The views expressed in this paper are mine and do not necessarily reflect those of the ECB or of the ESCB. Author contact: isabella.moder@ecb.europa.eu.

spillovers via real channels like trade links and remittance flows but, more and more, also the impact of financial spillovers, as monetary policy measures often generate sizable changes in capital flows and exchange rate dynamics. This mechanism could be very well observed in the so-called taper tantrum episode in mid-2013, when the Federal Reserve announced that it would gradually turn off its bond-buying program, which provoked a pronounced shift in market sentiment vis-à-vis emerging markets (see Sahay et al. 2014). Quantifying the direction and magnitude of international spillovers caused by advanced economies' monetary policy measures—and identifying the main transmission channels—is thus of utmost importance for policymakers in order to design optimum policy responses, both to spillovers from the introduction of such measures and to spillovers from their potential reversal.

The focus of interest for this paper lies in potential spillovers of ECB monetary policy measures to European countries that are not yet part of the euro area or are in the process of European Union (EU) accession. More specifically, this paper deals with the countries of southeastern Europe (SEE) that can be regarded as transition countries with respect to their economic development stage.[1] SEE countries are interlinked with the euro area through various channels. High trade integration and sizable remittance flows constitute potential real transmission channels, while the presence of a number of bank subsidiaries headquartered in the euro area and (correspondingly) a high degree of euroization[2] represent financial links.

Additionally, the heterogeneous monetary policy regimes of SEE countries provide an interesting case for cross-country comparisons with regard to the role of exchange rate regimes in shaping spillovers: exchange rate regimes in SEE range from inflation targeters with (managed) floating exchange rates (Albania, Romania, and Serbia), to stabilized arrangements with the euro as a reference currency (Croatia and North Macedonia), to euro-based

[1]This is in contrast to Baltic and central European countries, where convergence towards the "old" EU member states progressed further than in SEE countries.

[2]Either through high unofficial asset and liability euroization of the banking systems or, in the case of Montenegro, through the use of the euro as the legal tender; see European Central Bank (2016) for more information.

currency boards (Bosnia and Herzegovina as well as Bulgaria), to the unilateral adoption of the euro as the sole legal tender (Montenegro).

The aim of this paper is thus to answer three questions: First, in what direction and to which magnitude have the ECB's non-standard monetary policy measures been affecting the SEE countries? Second, through which channels are these shocks transmitted to SEE? Third, do different exchange rate regimes play a role in shaping the SEE countries' responses to a non-standard monetary policy shock?

The main contribution of this paper to the literature is the systematic examination of spillovers from the ECB's non-standard monetary policy measures to the whole SEE region. While its three EU members (Bulgaria, Croatia, and Romania) have already been covered to a certain extent in the spillover literature, no research has been undertaken yet for the remaining countries, which are five candidate and potential candidate countries to the EU.[3] By employing impulse response functions in a structural BVAR setting, the effect of non-standard monetary policy shocks on each country's output, price level, exports, short-term interest rate, and (if applicable) exchange rate is estimated.

The results show that the price level of all countries is positively affected by an expansionary non-standard monetary policy shock originating in the euro area, and for approximately half of the countries output also responds in a positive and pronounced way. Furthermore, spillovers seem to be mostly transmitted via exports, while in only a few cases the interbank interest rate (used as a proxy for the financial channel) exhibits a clear response to the shock. Additionally, the results suggest that in countries operating under a flexible exchange rate regime, exchange rates did not react strongly to the non-standard monetary shock, which is in line with the relatively stable exchange rates during the sample period.

The structure of the paper is as follows: Section 2 provides an overview of the literature, while section 3 introduces the methodological approach. The corresponding results and potential

[3]Due to data limitations, the remaining prospective EU member, Kosovo, cannot be included in the empirical analysis.

transmission channels are discussed both from a cross-country perspective and for each country individually in section 4. Section 5 documents robustness tests undertaken in order to cement the results. Section 6 concludes.

2. Related Literature

A vast amount of literature on cross-border monetary policy spillovers has emerged in the past decades, which in the beginning focused mostly on spillovers between advanced economies. Canova (2005) was among the first ones to investigate monetary policy spillovers from an advanced economy (the United States) to emerging economies (eight countries in Latin America), followed by other papers modeling spillovers from U.S. monetary policy to Latin America, Canada, and Asian economies (see, e.g., Maćkowiak 2007), and from euro-area monetary policy to other European countries (both emerging and advanced; see, e.g., Maćkowiak 2006, Jarociński 2010, and Benkovskis et al. 2011). On SEE countries, the literature on conventional monetary policy spillovers is less abundant, which is mainly related to the short time series available. Nevertheless, available results are very heterogeneous and seem to depend on the model and specifications used (see Jiménez-Rodriguez, Morales-Zumaquero, and Ègert 2010; Minea and Rault 2011; Feldkircher 2015; Petrevski et al. 2015; Hájek and Horváth 2016; and Potjagailo 2017, using near-VAR, VAR, GVAR, SVAR, and FAVAR models). Moreover, to the best of my knowledge, four SEE countries[4] have not yet been covered in the spillover literature at all.

The introduction of non-standard measures in October 2008 and the subsequent expansion and ongoing usage of several different unconventional instruments brought a new angle into the academic and policy discussion of euro-area monetary policy spillovers to countries outside the euro area.[5] However, given that the global experience with non-standard monetary policy is restricted (with a

[4]Namely Albania, Bosnia and Herzegovina, Montenegro, and Serbia.

[5]For the purpose of this paper, only spillovers outside the euro area are discussed. For an assessment of the effects *within* the euro area see, e.g., Peersman (2011); Boeckx, Dossche, and Peersman (2017); Burriel and Galesi (2018).

few exceptions) to the aftermath of the global financial crisis, the literature on spillovers from advanced economies' non-standard or unconventional monetary policy measures to emerging markets is relatively scarce. It can be divided into two categories: One strand investigates the impact of monetary policy *announcements* (and in some cases also actions) on high-frequency financial indicators (e.g., sovereign bond yields, stock market indexes, CDS spreads, or exchange rates); see, for example, Fratzscher, Lo Duca, and Straub (2016) for spillovers of U.S., and Georgiadis and Gräb (2015) as well as Falagiarda, McQuade, and Tirpák (2015) for spillovers of euro-area non-standard monetary policy. The latter examine the effects of more than seventy announcement-related events on financial assets of four non-euro-area EU countries. For Romania, which is the only SEE country covered, they find a significant effect on the short-term money market rate and an especially pronounced effect on long-term government bond yields, while the exchange rate seems not to respond immediately to an ECB announcement. Ciarlone and Colabella (2016) test the effect of the ECB's asset purchase programs on a panel of eleven countries in central, eastern, and southeastern Europe, including all countries covered in this paper. They find significant short-term spillover effects on financial variables as well as long-term spillovers on portfolio and cross-border banking flows.

Focusing on longer-lasting macroeconomic effects instead, the second strand of literature has been following the methods of the literature on "conventional" monetary policy spillovers by using some kind of VAR model to assess spillovers on macroeconomic variables. The literature on non-standard monetary policy spillovers from the euro area to central and eastern Europe (CEE) is scarce, whereas it is non-existent for most countries in SEE. Babecká Kucharčuková, Claeys, and Vašíček (2016) investigate spillovers on six EU non-euro-area countries, among them three in CEE (Czech Republic, Hungary, and Poland). They conclude that the spillovers of unconventional shocks are transmitted differently than those of conventional shocks, and while exchange rates respond quickly, the effect on inflation is ambiguous. Bluwstein and Canova (2016) use a Bayesian mixed-frequency VAR to incorporate both high-frequency financial and low-frequency macroeconomic data. They find that output effects of unconventional monetary policy measures were insignificant for

CEE countries (Czech Republic, Hungary, and Poland) and slightly negative for SEE countries (Bulgaria and Romania), and that the impact on inflation was slightly positive for both groups. With regard to the exchange rate channel, they conclude that it does not seem to shape the response of macroeconomic variables in the case of unconventional monetary policy shocks, as opposed to the case for conventional monetary policy. Halova and Horváth (2015) employ a PVAR model for eleven CEE and SEE countries (among them Bulgaria, Croatia, and Romania). Contrary to Bluwstein and Canova (2016) and Babecká Kucharčuková, Claeys, and Vašíček (2016), their results suggest sizable spillovers and that a significant amount of output fluctuations in the CEE and SEE countries can be explained by the euro area's non-standard monetary policy measures. Ultimately, whether a country benefits from or is negatively affected by spillovers of a foreign monetary policy shock depends on whether its business cycle is in the same position as that of the "core country" (Chen et al. 2015).

Another open issue that has been discussed in the spillover literature is what shapes the response of an economy to a foreign monetary policy shock (both conventional and non-standard). The role of the exchange rate regime has featured prominently in the spillover discussion, following the argument that flexible exchange rates are better suited to buffer real external shocks (based on Meade 1951 and Friedman 1953). Other potential determinants identified by the literature are the degrees of trade and financial openness (see, e.g., Miniane and Rogers 2007). More recently, Georgiadis (2016) systematically examines U.S. monetary policy spillovers and finds that the role of the exchange rate regime is non-linear and that non-advanced economies operating under an inflexible exchange rate regime experience larger spillovers the more strongly they are integrated in global trade. Furthermore, the results suggest that trade integration amplifies spillovers to non-advanced economies if the share of manufactured goods in aggregate output is large and the country participates in global value chains. Crespo Cuaresma et al. (2016) additionally find that macroeconomic vulnerabilities such as high external imbalances tend to amplify spillovers from U.S. monetary policy shocks. For spillovers of euro-area monetary policy, Potjagailo (2017) presents some evidence that spillovers on other EU countries' output are larger if the exchange rate regime is fixed.

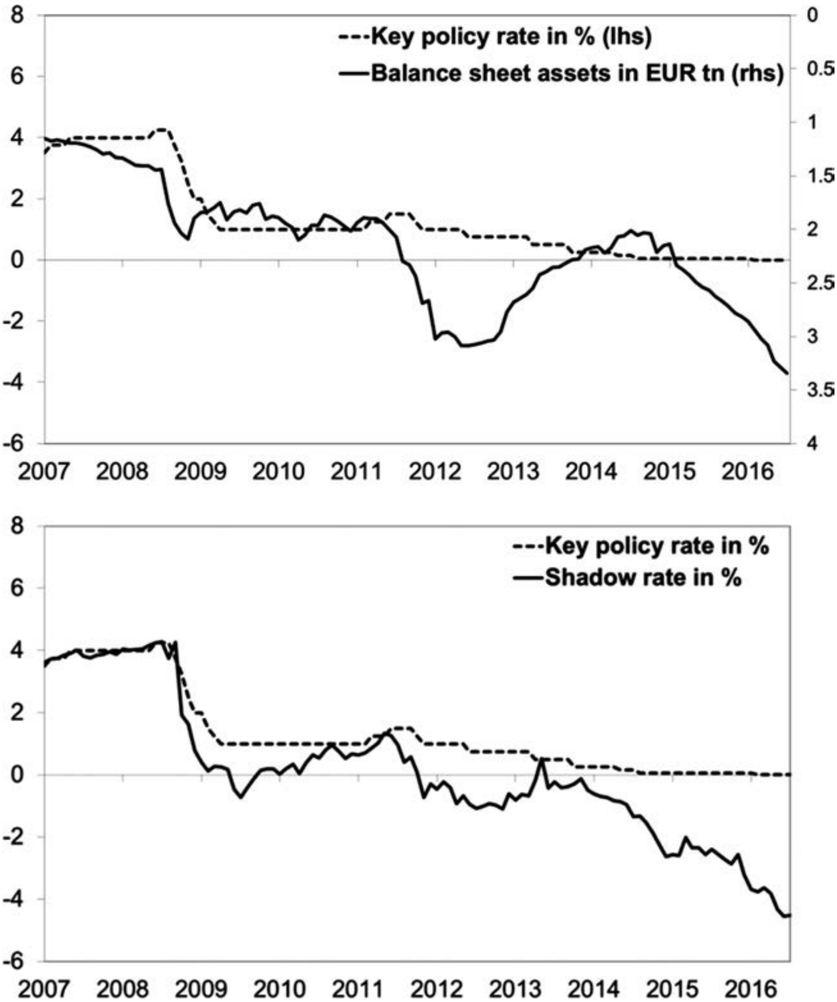
3. Methodology

The methodology used in this paper follows the strand of literature that investigates the effects of non-standard monetary policy measures on the real economy by employing some specification of a VAR model. There are two reasons to choose this approach: First, although event studies could identify significant financial market spillovers for some European countries outside the euro area, this does not necessarily imply that the real economy is equally affected, since financial variables often exhibit overshooting behavior that does not necessarily transmit into the real economy. Second, the event-study approach requires developed financial markets to investigate the behavior of high-frequency indicators. This is a major drawback in the case of emerging markets in general and SEE in particular, as these countries have very shallow financial markets in line with their small economic size and comparatively low GDP per capita levels.

3.1 Issues in Dealing with Non-standard Monetary Policy

Empirically assessing non-standard monetary policy measures brings a number of additional challenges compared with conventional monetary policy. First, as the key monetary policy rate does not incorporate non-standard measures, alternative indicators for the stance of non-standard monetary policy have to be found. Those used in the literature so far have been the term spread between government bonds of different maturities (e.g., Chen et al. 2012), central bank balance sheet assets (e.g., Gambacorta, Hofmann, and Peersman 2014), or shadow rates that are supposed to be directly comparable to key policy rates (Lombardi and Zhu 2014; Krippner 2015; Wu and Xia 2016). In this paper I use Eurosystem balance sheet assets as the main measure of non-standard monetary policy. In the upper chart of figure 1, inverted Eurosystem balance sheet assets are plotted together with the key policy rate. It can be seen that the key policy rate was decreased in various steps to 1 percent in May 2009, from where it slowly and gradually moved towards the zero lower bound, which was reached in March 2016. In contrast, the Eurosystem balance sheet assets started to increase already with the switch of liquidity operations to a fixed-rate tender with full allotment in

Figure 1. Indicators of Non-standard Monetary Policy Measures



Sources: ECB, Wu and Xia (2016).

October 2008, and thereafter fluctuated with the introduction and phase-out of the different programs. In this paper I use additionally the shadow rate developed by Wu and Xia (2016) for robustness testing. It is calculated by assessing bond prices in a framework of a multifactor term structure model and is directly comparable to the key policy rate, as both interest rates are equal in conventional times (see the lower chart of figure 1). In contrast to balance sheet assets, this indicator also includes announcement effects of non-standard monetary policy measures whenever they affect bond yields.

Second, the way some of the ECB's non-standard monetary policy measures have been designed makes it necessary to find an empirical strategy that disentangles exogenous monetary policy shocks from endogenous or demand-driven monetary expansions. Since the change in the operational framework from standard tender-based allotment to fixed-rate full allotment in October 2008, monetary policy operations have been essentially endogenous or demand driven, as banks have unlimited access to liquidity at the interest rate on the main refinancing operations (MRO) under the condition that they can provide enough collateral (Boeckx, Dossche, and Peersman 2017). Moreover, the (targeted) longer-term refinancing operations, which were increased in duration and size in October 2008, are also endogenous to a certain extent since the ECB only fixes the upper ceiling of these operations, whereas banks decide how much to draw upon that limit. This paper deals with endogeneity issues in several ways. First, I follow the approach proposed by Boeckx, Dossche, and Peersman (2017), who complement the Eurosystem's balance sheet assets as the main measure for non-standard monetary policy measures with certain assumptions on shock identification (see subsection 3.3). Moreover, I perform robustness checks by using only the position "Securities held for monetary policy purposes" (A070100) of the Eurosystem's balance sheet, which incorporates all securities purchased under the various purchasing programs. Compared with other positions of the Eurosystem balance sheet, this is the most exogenous part, since the size and frequency of bond purchases are ex ante determined by the ECB and not shaped by banks' behavior.

3.2 Model

To model spillovers from the euro area's non-standard monetary policy to SEE countries, the following structural BVAR model with a monthly frequency is employed for each SEE country:

$$\sum_{s=0}^{p} \begin{bmatrix} A_{11}(s) & A_{12}(s) \\ A_{21}(s) & A_{22}(s) \end{bmatrix} \begin{bmatrix} y_1(t-s) \\ y_2(t-s) \end{bmatrix} + \begin{bmatrix} c_{11} \\ c_{21} \end{bmatrix} = \begin{bmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{bmatrix},$$

where $y_1(t)$ represents a vector of macroeconomic variables of the SEE country, $y_2(t)$ represents a vector of macroeconomic variables of the euro area, and the vectors c_{11}, c_{21} are constants. The vectors $\varepsilon_1(t) \sim N(0, \Sigma_1)$ and $\varepsilon_2(t) \sim N(0, \Sigma_2)$ denote structural shocks of domestic and euro-area origin, respectively.

For each s, $A_{21}(s) = 0$, implying that the variables of the SEE country are set to be exogenous to the variables of the euro area under the assumption that neither current nor past economic developments in the SEE countries influence developments in the euro area. This so-called block exogeneity feature introduced by Cushman and Zha (1997) has been used frequently in the literature (see, e.g., Canova 2005; Maćkowiak 2007; Benkovskis et al. 2011) and is well suited for modeling spillovers from large to small economies, as it helps to identify spillovers from the viewpoint of the small open economy and reduces the number of parameters to be estimated (Cushman and Zha 1997).

The vector y_1 consists of the following variables:

$$y_1 = \begin{pmatrix} y_t^{SEE} & p_t^{SEE} \end{pmatrix}',$$

where y_t^{SEE} denotes output and p_t^{SEE} denotes prices of the respective SEE country. At a second stage, in order to investigate potential transmission channels, the vector y_1 includes either exports,[6] x_t^{SEE}, the interbank market rate of the respective SEE country, i_t^{SEE}, or the exchange rate of the local currency vis-à-vis the euro, e_t, for countries that are operating under a flexible exchange rate regime. The vector y_2 represents the euro area and includes six variables:

$$y_2 = \begin{pmatrix} y_t^{EA} & p_t^{EA} & assets_t^{EA} & CISS_t & spread_t^{EA} & MRO_t \end{pmatrix}',$$

where y_t^{EA} and p_t^{EA} again denote output and prices, respectively, but this time for the euro area, and $assets_t^{EA}$ represents Eurosystem balance sheet assets as the main measure for non-standard monetary policy (as discussed in subsection 3.1). Moreover, following Gambacorta, Hofmann, and Peersman (2014), Boeckx, Dossche, and Peersman (2017), and Burriel and Galesi (2018), I include the

[6] In order to account for indirect spillovers, exports to the world (instead of only to the euro area) are used.

CISS indicator (composite indicator of systemic stress) developed by Holló, Kremer, and Lo Duca (2012) ($CISS_t$), which serves two purposes: First, it controls for the impact of euro-area financial stress and economic risk, which is important to capture in the model, as it has had pronounced effects on euro-area macroeconomic developments. Second, the inclusion of the CISS indicator helps to disentangle exogenous balance sheet movements from endogenous ones and thus enables a proper identification of monetary policy shocks (see subsection 3.3). For the same purpose, I also include the spread between EONIA and the MRO rate (denoted $spread_t^{EA}$). To disentangle conventional from non-standard monetary policy shocks, the model incorporates additionally the MRO rate (MRO_t).

The chosen estimation procedure is Bayesian, because it is better suited for shorter data sets compared with frequentist methods. I use an independent normal-Wishart prior and obtain the scale matrix S_0 from individual AR regressions. Estimations are carried out by employing the BEAR (Bayesian estimation, analysis, and regression) toolbox developed by Dieppe, Legrand, and van Roye (2016). The autoregressive coefficient of the prior is set to 1, since the variables enter the model in levels. This specification is possible in Bayesian models, where the prior can account for unit-root behavior by including an autoregresssive coefficient on the first own lag of each variable, and it has the advantage of avoiding the transformation bias that occurs when data enter transformed into first differences. The remaining hyperparameters that specify the prior are chosen following Dieppe, Legrand, and van Roye (2016).

The posterior is derived by Gibbs sampling, with a total number of 5,000 iterations and a burn-in sample of 1,000 iterations. The Bayesian information criterion (BIC) suggests a lag length of 1; however, testing for autoregressive behavior of the residuals suggests that a model specification of four lags is best to avoid residual autocorrelation. Therefore I define $p = 4$.

3.3 Identification

In order to generate impulse response functions, the identification of shocks is carried out via sign and zero restrictions, following the method proposed by Arias, Rubio-Ramirez, and Waggoner (2014) (see Dieppe, Legrand, and van Roye 2016). The non-standard

Table 1. Sign and Zero Restrictions for the Shock Identification of the Baseline Model

$assets_t^{EA}$	$CISS_t$	$spread_t^{EA}$	MRO_t	y_t^{EA}	p_t^{EA}	y_t^{SEE}	p_t^{SEE}
+ 0–1	– 0–1	– 0–1	0	0	0	0	0

Note: 0 indicates that the immediate response is restricted, while + (–) indicates that only a positive (negative) reaction is permitted in the respective period.

monetary policy shock is the only identified shock in the model (see table 1). The first six variables define the non-standard monetary policy shock and its effects on the euro area, while the remaining variables apply to the respective SEE country's output and price level. An expansionary non-standard monetary policy shock increases the Eurosystem balance sheet assets on impact and in the first month following the shock, while the CISS indicator as well as the spread between EONIA and the MRO decrease immediately (on impact) and in the first month after the shock. These identifying assumptions are taken to distinguish demand-driven from exogenous balance sheet shocks, following Boeckx, Dossche, and Peersman (2017) and Burriel and Galesi (2018). More specifically, in periods of financial stress or other shocks, increased demand for liquidity expands the balance sheet, implying that the CISS indicator as well as EONIA increase (see Boeckx, Dossche, and Peersman 2017). Vice versa, a balance sheet expansion that is caused by an ECB monetary policy measure should not increase but *decrease* both financial stress and the demand for liquidity, which is exactly the assumption taken in the shock identification. Finally, the zero restriction of the MRO rate ensures that the balance sheet increase is orthogonal to a conventional monetary policy shock. For the response of output and prices, I follow the standard approach of defining conventional monetary policy shocks by imposing zero restrictions to disentangle it from other shocks. Similarly, zero restrictions are placed on output and price responses of the SEE country, in order to disentangle the potential spillover from domestic real economy disturbances.[7] The

[7]For the same reason, a zero restriction is put on exports in the subsequent estimations on potential transmission channels.

Table 2. Acceptance Rates of Structural Matrices (in %)

Albania	13.93
Bosnia and Herzegovina	12.21
Bulgaria	13.71
Croatia	11.95
North Macedonia	13.40
Montenegro	12.24
Romania	11.96
Serbia	13.12

non-standard monetary policy shock is the only shock identified. The acceptance rates of the structural matrices from the baseline models are depicted in table 2.

3.4 Data

The time span of the baseline model covers the period between January 2008 and December 2017. As a measure of output I use GDP, interpolated by the Chow-Lin method with industrial production to obtain data with monthly frequency.[8] The price level is measured as the consumer price inflation (harmonized consumer price inflation in the case of EU countries) index. As indicator for a potential financial channel I use short-term interest rates, since time series on asset prices or longer-term interest rates are not available for all countries. More specifically, I include monthly values of three-month interbank market rates, with the exception of Bosnia and Herzegovina as well as Montenegro, which do not publish interbank market rates. In the case of Montenegro, an unweighted average of three-month and six-month government T-bill rates is used as a proxy for interbank market rates. For Bosnia and Herzegovina, no such short-term interest rate exists. Therefore, following the approach of Cerutti et al. (2010), I create a composite series that consists of two-thirds of retail deposit rates and one-third of retail lending

[8] The exception is Montenegro, where GDP is not available for the whole time span and therefore industrial production is used as a measure for output.

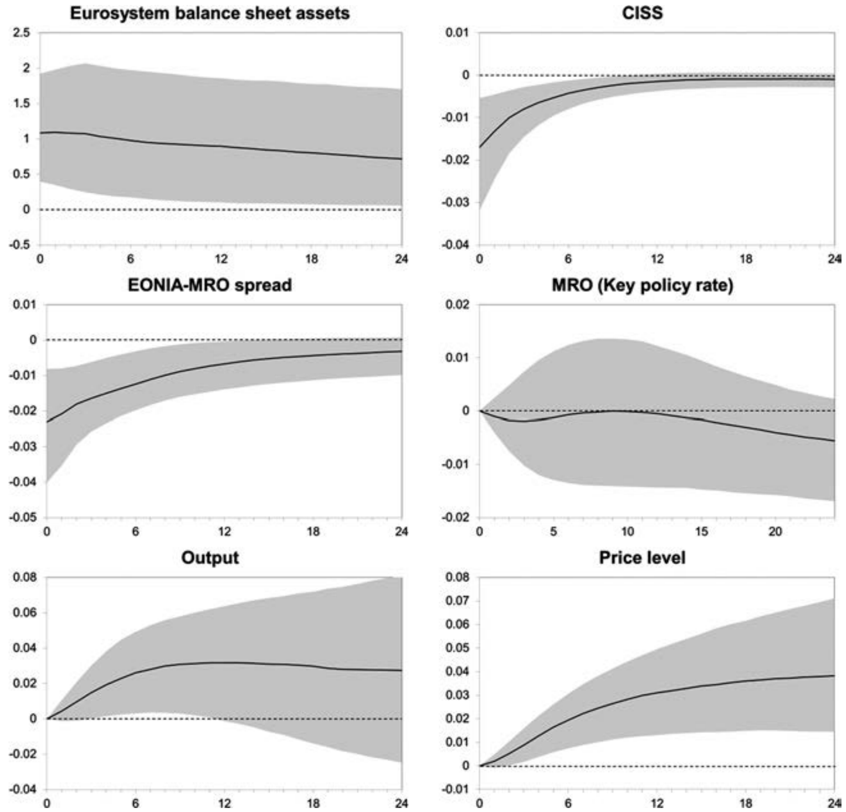
rates, both in the corporate sector. All exchange rates are expressed in average local currency vis-à-vis the euro, so an increase in the exchange rate depicts depreciation and vice versa. Data on exports, which are derived from the International Monetary Fund's (IMF's) Direction of Trade Statistics and converted into euros, are limited to merchandise exports, which means that service exports are not captured in the model. All variables enter the model in monthly frequency and in levels. Moreover, all variables are seasonally adjusted by the U.S. Census Bureau's X-13 seasonal adjustment procedure and are transformed into their natural logs (with the exception of financial variables). The data sources are national central banks, national statistical offices, Eurostat, the ECB, the IMF's Direction of Trade Statistics, and Bloomberg.

4. Results

4.1 The Effect of a Non-standard Monetary Policy Shock on the Euro-Area Economy

I start by looking at the transmission of an expansionary Eurosystem balance sheet shock within the euro area. The impulse response functions of the euro area are displayed in figure 2, where the continuous line depicts the median posterior response and the shaded area represents 68 percent of the credibility interval. Because the variables enter the model in natural logs, the y-axis reports percentage changes, except for financial variables, which are depicted in percentage-point changes. It can be observed that the one-standard-deviation (1.1 percent increase) balance sheet shock is very persistent, as it remains almost unchanged for two years. Both the accompanying decline of the CISS indicator and the decrease of the EONIA-MRO spread are less persistent but do not fade out completely until the end of the two-year horizon. The response of the MRO (key policy rate) to the non-standard monetary policy shock is ambiguous. Turning to the macroeconomic effects, the impulse responses suggest that output rises gradually, with a peak increase of 0.03 percent after eleven months. The price-level increase reaches 0.04 percent after two years and, in line with economic theory, the response seems to be relatively persistent.

Figure 2. Euro Area: Response to a Balance Sheet Shock

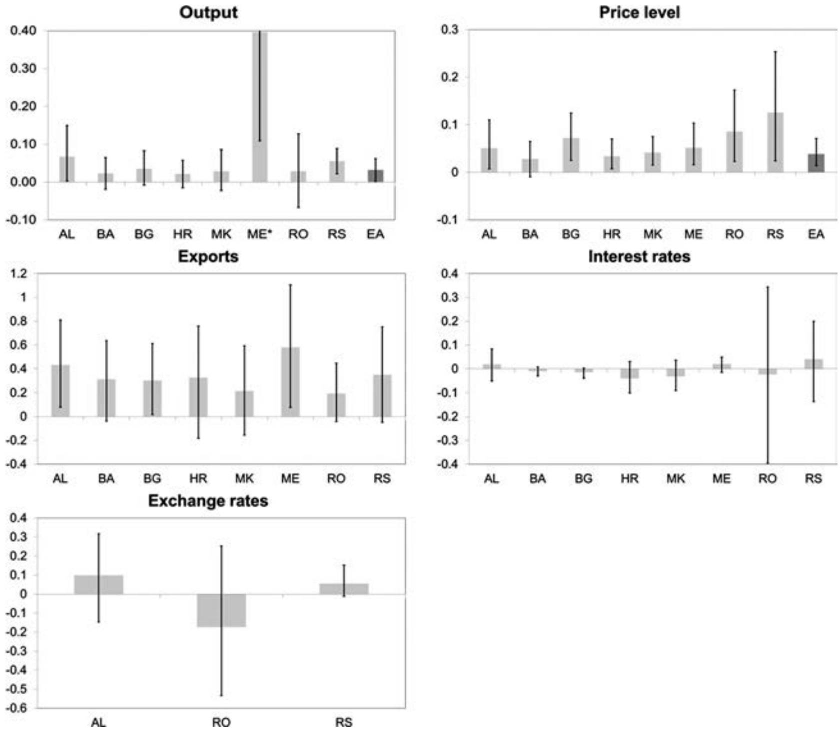


Notes: The figure shows the response of variables to an expansionary one-standard-deviation Eurosystem balance sheet shock. The shaded regions report pointwise 68 percent credibility intervals. The x-axis reports months, and the y-axis reports monthly growth rates in percent for all variables except for financial variables, where percentage-point changes are depicted.

4.2 Spillovers to SEE Countries and Their Transmission Channels

Turning to the SEE countries, the results suggest that spillovers from a euro-area expansionary non-standard monetary policy shock are positive (see figure 3, upper-left panel). In the three countries that exhibit the highest magnitude in their shock response

Figure 3. Comparison of Responses to a Non-standard Monetary Policy Shock



Note: The figure shows the peak response to an expansionary balance sheet shock within the first twenty-four months in percent.
*For Montenegro, output refers to industrial production instead of GDP. The results are derived from the baseline model (output and price level) as well as separate models including exports, interest rates, and exchange rates.

(Montenegro,[9] Albania, and Serbia), it is credible within the 68 percent interval. Moreover, in half of the countries the response of output seems to be stronger than in the euro area.

Price-level responses are positive in all countries and lie (with one exception) within the 68 percent credibility interval (see figure 3, upper-right panel). It can be observed that the peak response is by

[9]It should be noted that for Montenegro, the response of industrial production instead of GDP is depicted, which explains the magnitude.

far the strongest for Serbia (at 0.13 percent), followed by Romania and Bulgaria. The relatively strong price responses, which are in most countries stronger than the euro-area response itself, are in line with the high share of imports from the euro area that range from around one-third to over 50 percent of all imports in SEE countries.

In order to shed light on potential transmission channels, I estimate spillovers on exports and interbank interest rates in separate models, where the vector $y_2(t)$ for the identification of the euro-area shock remains unchanged but the vector $y_1(t)$ contains either exports or short-term interest rates. For exports, the peak response is depicted in the center-left panel of figure 3. The impact of a non-standard monetary policy shock on SEE exports is positive in all countries, suggesting that exports are indeed an important channel of shock transmission, and it is credible within a 68 percent interval in three countries. The largest magnitude of the shock response can be observed for Montenegro, Albania, and Serbia, which corresponds to the relative magnitude of their output reaction and thus suggests that exports are indeed a relevant transmission channel.

Turning to short-term interest rates, the peak responses across countries are heterogeneous in sign and magnitude, and surrounded by large uncertainty bands. One reason for the weak model output might be the relatively illiquid interbank money markets in SEE countries. Thus this result should not necessarily be taken as proof that financial channels do not transmit non-standard monetary policy shocks from the euro area to SEE countries, as changes in the interbank market rate do not capture foreign direct or portfolio inflows.

Comparing the peak responses across countries, the results suggest that the inflation-targeting countries which operate under a managed or flexible exchange regime were equally affected by spillovers as countries that have pegged their currency to the euro. This result is not surprising when the model output on exchange rate responses is taken into account (figure 3, lower-left panel), where the peak responses in Albania and Serbia suggest an exchange rate depreciation (rather than an appreciation, which would be expected by economic theory) and in Romania the peak appreciation is surrounded by a large uncertainty band. This is in line with the very stable exchange rates observed for Albania and Romania in the sample period. Moreover, between January 2008 and December 2017 the

Table 3. Forecast Error Variance

	Output	Price Level
Euro Area	1.24	4.94
Albania	3.98	3.10
Bosnia and Herzegovina	0.67	1.21
Bulgaria	1.45	4.23
Croatia	0.73	1.66
North Macedonia	1.04	2.68
Montenegro	1.29	3.78
Romania	0.83	4.27
Serbia	2.45	3.04

Note: Percentage of the variance of the respective variables explained by a non-standard monetary policy shock after twenty-four months.

central banks of Albania, Romania, and Serbia eased their monetary policy stance by decreasing their interest rates by a total of 500, 625, and 650 basis points, respectively, which might have counteracted appreciation pressures on the exchange rate. The finding that in SEE flexible exchange rates did not respond strongly to non-standard monetary policy shocks during the period under review is in line with Bluwstein and Canova (2016), who argue that the exchange rate channel is not important for unconventional monetary policy transmission (which is different from the conventional case).

4.3 The Importance of Euro-Area Non-standard Monetary Policy Shocks for SEE Countries

Besides analyzing the peak effect from non-standard monetary policy shocks on certain macroeconomic variables in SEE, it is also important to assess how much of the variance of output and prices can be explained by non-standard monetary policy shocks. Table 3 presents the percentage of the forecast error variance of output and price explained by a non-standard monetary policy shock. For the euro area itself, non-standard monetary policy shocks account for

1.24 percent of the fluctuations of GDP and 4.94 percent of the fluctuations of the price level after two years. In a similar vein, for most SEE countries the degree of variance explained by euro-area non-standard monetary policy shocks is higher for prices than for output. Comparing the forecast error variance of output among SEE countries, Albania, Serbia, and Bulgaria are the countries most exposed to euro-area non-standard monetary policy shocks, and in half of the countries the movements in output are larger than the output variability in the euro area, which is in line with the analysis from impulse response functions. With regard to the price level, the highest variability among SEE countries can be observed in Romania, Bulgaria, and Montenegro, which is again mostly in line with the previous analysis. However, the explanatory power of the non-standard monetary policy shock in price levels is lower in SEE countries than in the euro area.

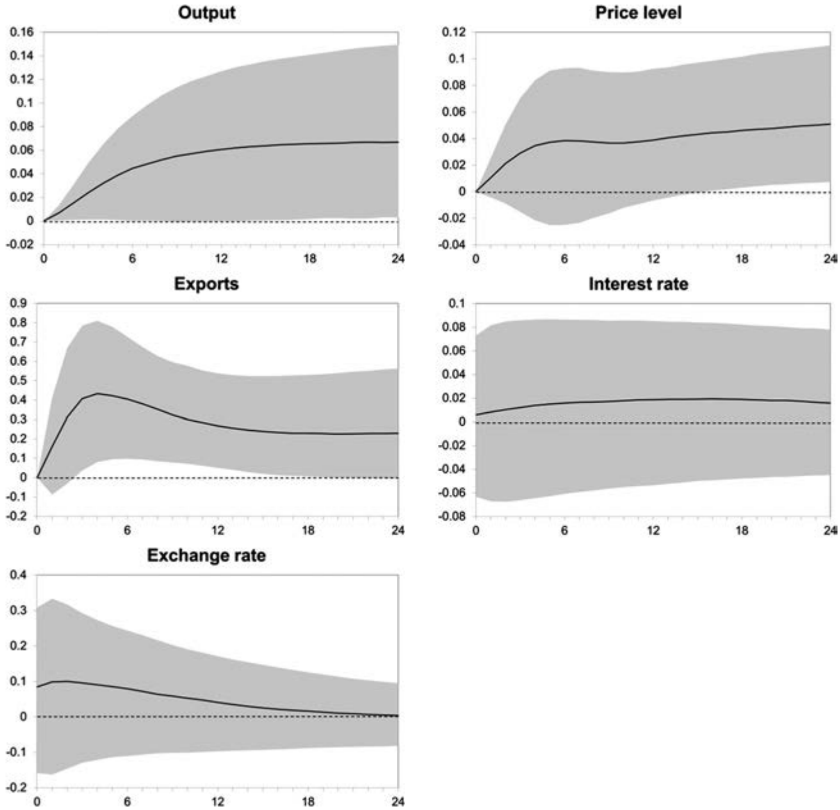
4.4 Results for Individual SEE Countries

The response of the Albanian economy to a non-standard monetary policy shock is shown in figure 4. An expansionary euro-area balance sheet shock raises output by 0.07 percent and prices by 0.05 percent after two years. Albania's exports rise as a response to the balance sheet shock, peaking at 0.43 percent after five months, suggesting that exports are an important transmission channel to explain the relatively strong spillover to the Albanian economy. The response of the interbank interest rate on the other hand seems to be muted. Also, the exchange rate of the lek vis-à-vis the euro does not exhibit an unambiguous response to the shock, which is in line with the fact that it has fluctuated only slightly against the euro in the sample period.[10]

In the case of Bosnia and Herzegovina, the economic response to a non-standard euro-area monetary policy shock is depicted in figure 5. Output shows initially a slightly negative response which turns positive after seven months, but the uncertainty band surrounding the response is relatively high. Notwithstanding the mixed output response, the price level increases gradually, with a peak

[10]Between 2008 and 2017 the average monthly fluctuation against the euro amounted to 0.5 percent.

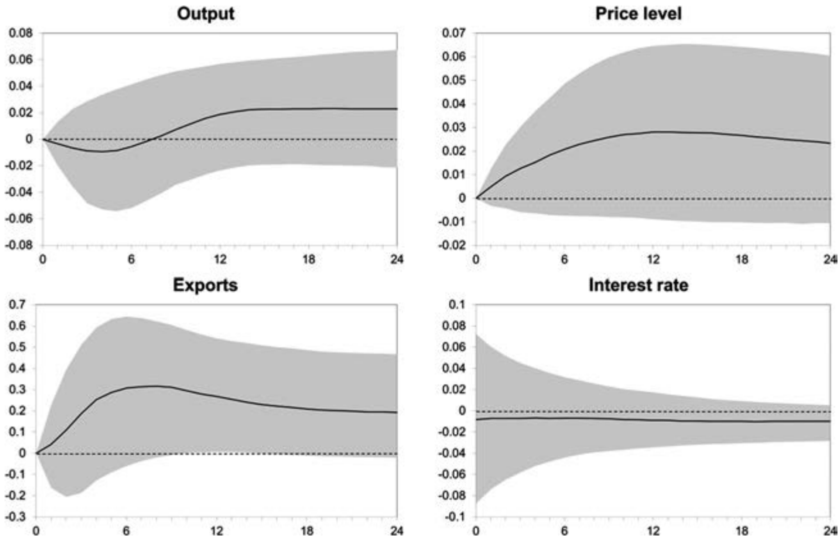
Figure 4. Albania: Response to a Balance Sheet Shock



Notes: The figure shows the response of variables to an expansionary one-standard-deviation Eurosystem balance sheet shock. The shaded regions report pointwise 68 percent credibility intervals. The x-axis reports months, and the y-axis reports monthly growth rates in percent for all variables except for interest rates, which depict changes in percentage points. The results are derived from the baseline model (output and price level) as well as separate models including exports, the interest rate, and the exchange rate.

increase of 0.03 percent after twelve months. Exports react in a pronounced manner, with a peak response of 0.32 percent after nine months. The uncertainty band of the interest rate response (which in Bosnia and Herzegovina's case is a composite retail rate) is relatively wide in the short term but becomes more significant in the medium term, with a decrease of 0.01 percentage points.

Figure 5. Bosnia and Herzegovina: Response to a Balance Sheet Shock

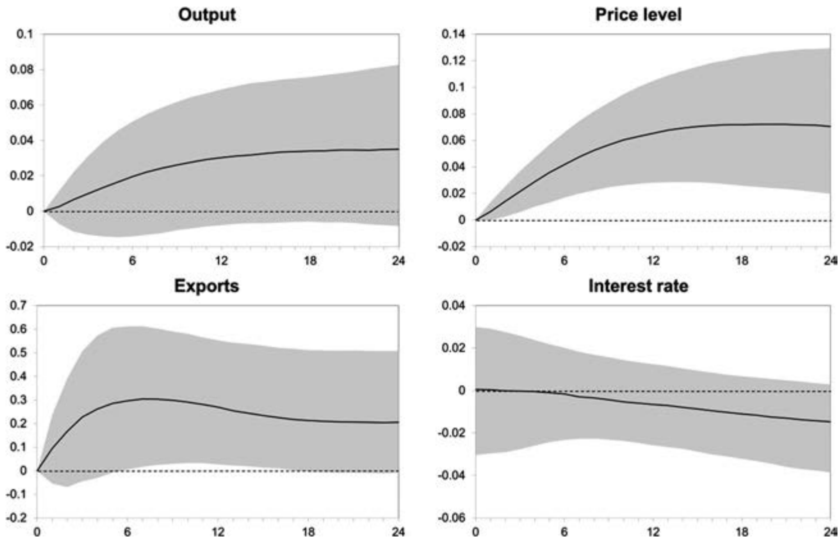


Notes: The figure shows the response of variables to an expansionary one-standard-deviation Eurosystem balance sheet shock. The shaded regions report pointwise 68 percent credibility intervals. The x-axis reports months, and the y-axis reports monthly growth rates in percent for all variables except for the interest rate, which depicts changes in percentage points. The results are derived from the baseline model (output and price level) as well as separate models including exports and the interest rate.

For Bulgaria (see figure 6), the output response is positive, with an increase of 0.04 percent after two years. The price level exhibits a pronounced increase, reaching the peak of 0.07 percent after nineteen months. The export channel seems to be relevant also in the case of Bulgaria, as exports rise, with a peak of 0.30 percent after eight months. The interbank interest rate does not react in the short run, but in the medium term it exhibits a decrease of 0.01 percentage point at the end of the two-year horizon. The marked reaction of output and prices in Bulgaria are in line with the results of Hájek and Horváth (2016) for spillovers of positive short-term interest rate shocks.

In the case of Croatia (see figure 7), the response of output is positive and peaks at 0.02 percent after twelve months, although it

Figure 6. Bulgaria: Response to a Balance Sheet Shock

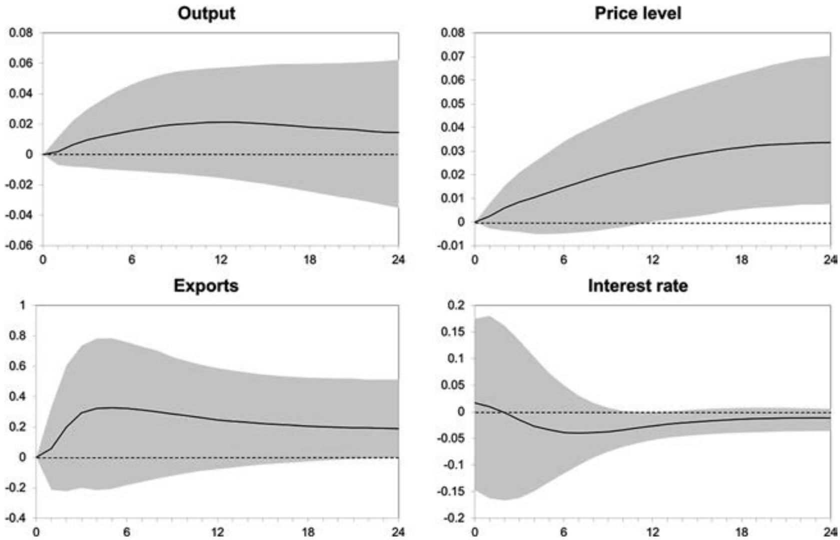


Notes: The figure shows the response of variables to an expansionary one-standard-deviation Eurosystem balance sheet shock. The shaded regions report pointwise 68 percent credibility intervals. The x-axis reports months, and the y-axis reports monthly growth rates in percent for all variables except for the interest rate, which depicts changes in percentage points. The results are derived from the baseline model (output and price level) as well as separate models including exports and the interest rate.

is surrounded by some uncertainty. The response of the price level is more significant and exhibits an increase of 0.03 percent at the end of the two-year horizon. Both the output and the price response are in line with what Hájek and Horváth (2016) find for policy spillovers of positive short-term interest rate shocks.[11] Furthermore, the export response peaks at 0.33 percent after six months, while the interbank market rate decreases, with a trough of 0.04 percentage point after eight months. The results suggest that both exports and financial channels might be relevant in transmitting shocks from non-standard monetary policy measures in the euro area to Croatia.

[11] Conversely, Petrevski et al. (2015) find that the same shock *increases* Croatia's price level.

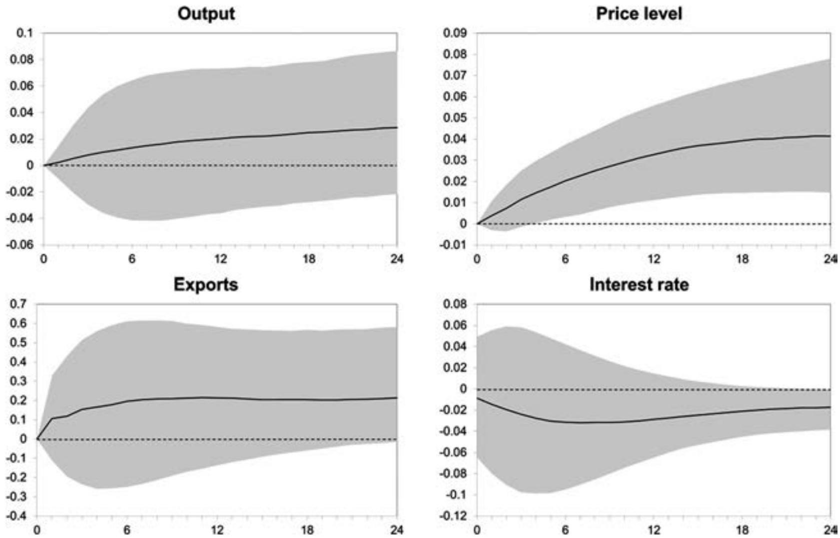
Figure 7. Croatia: Response to a Balance Sheet Shock



Notes: The figure shows the response of variables to an expansionary one-standard-deviation Eurosystem balance sheet shock. The shaded regions report pointwise 68 percent credibility intervals. The x-axis reports months, and the y-axis reports monthly growth rates in percent for all variables except for the interest rate, which depicts changes in percentage points. The results are derived from the baseline model (output and price level) as well as separate models including exports and the interest rate.

The response of the economy of North Macedonia is depicted in figure 8. The results suggest that the euro-area non-standard monetary policy shock does not trigger a significant output response, while the price-level response is positive, with a peak of 0.04 percent after twenty-three months. Exports peak—although not strictly significantly—at 0.22 percent after twelve months. The interbank market rate decreases by a maximum of 0.03 percentage point after eight months, with the response being relatively persistent and becoming more significant towards the end of the horizon, suggesting that financial spillovers could play a role at least in the medium term. The results for the interest rate response are in line with Petrevski et al. (2015) for a positive short-term interest rate shock; however, they find a different response of the price level.

Figure 8. North Macedonia: Response to a Balance Sheet Shock

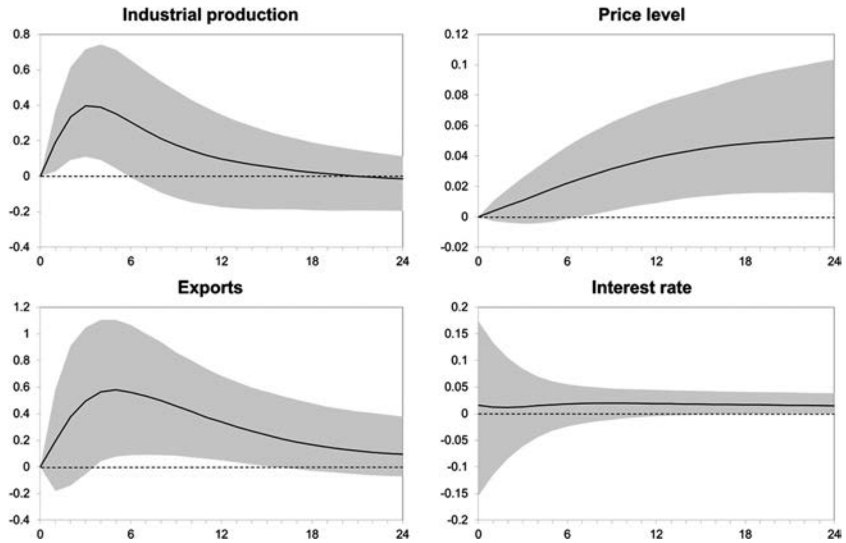


Notes: The figure shows the response of variables to an expansionary one-standard-deviation Eurosystem balance sheet shock. The shaded regions report pointwise 68 percent credibility intervals. The x-axis reports months, and the y-axis reports monthly growth rates in percent for all variables except for the interest rate, which depicts changes in percentage points. The results are derived from the baseline model (output and price level) as well as separate models including exports and the interest rate.

For Montenegro, an exogenous expansion of the Eurosystem's balance sheet translates into a pronounced rise of industrial production by 0.40 percent after three months (see figure 9). The price level also increases by 0.05 percent at the end of the two-year horizon. Montenegro's exports seem to rise by 0.58 percent after six months, suggesting that an increase in exports might explain the rise in industrial production and prices. The uncertainty of the interest rate response, which is in the case of Montenegro a composite of three- and six-month T-bill rates, is very high in the short term, while in the medium term the increase in interest rates becomes more significant.

Spillovers to Romanian output from the non-standard monetary policy shock are not very pronounced (compare figure 10). This

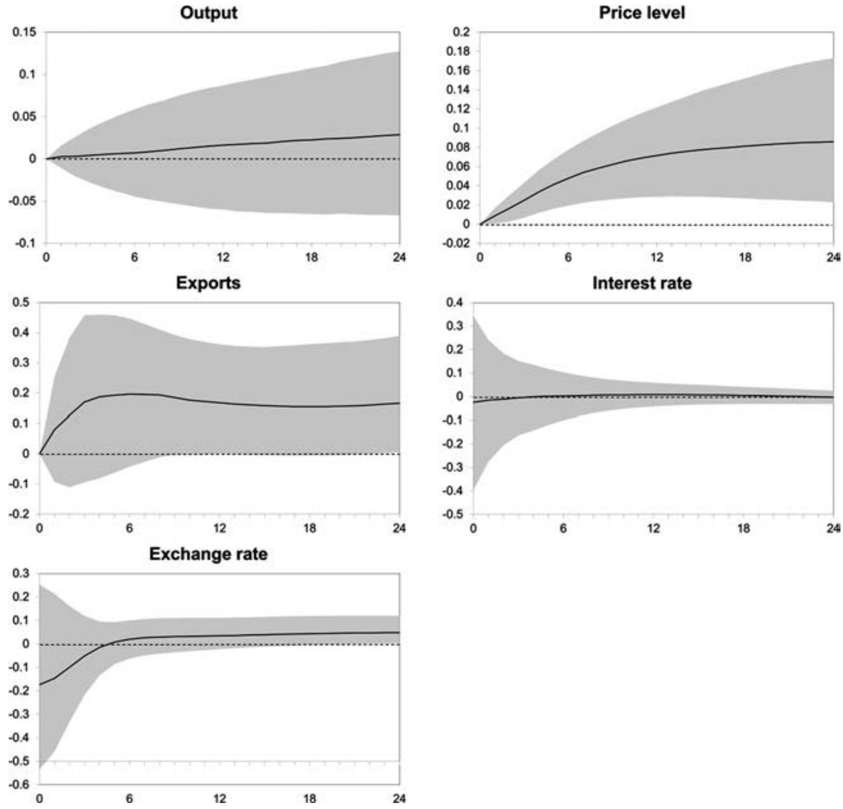
Figure 9. Montenegro: Response to a Balance Sheet Shock



Notes: The figure shows the response of variables to an expansionary one-standard-deviation Eurosystem balance sheet shock. The shaded regions report pointwise 68 percent credibility intervals. The x-axis reports months, and the y-axis reports monthly growth rates in percent for all variables except for the interest rate, which depicts changes in percentage points. The results are derived from the baseline model (output and price level) as well as separate models including exports and the interest rate.

result is different from the findings of Hájek and Horváth (2016) and Bluwstein and Canova (2016), who conclude that a contractionary euro-area short-term interest rate shock initially increases Romanian output (or that an expansionary non-standard monetary policy output decreases Romanian output, respectively). The response of prices, on the other hand, is positive and relatively strong, with an increase of 0.09 percent after two years. Exports are peaking at 0.20 percent after a period of seven months. On the contrary, the response of the short-term interest rate is muted, suggesting that the euro-area shock does not affect the interbank market rate in Romania. Initially the exchange rate seems to appreciate, which is however subject to high uncertainty, which turns into a more pronounced and persistent depreciation after five months. The blurred response

Figure 10. Romania: Response to a Balance Sheet Shock

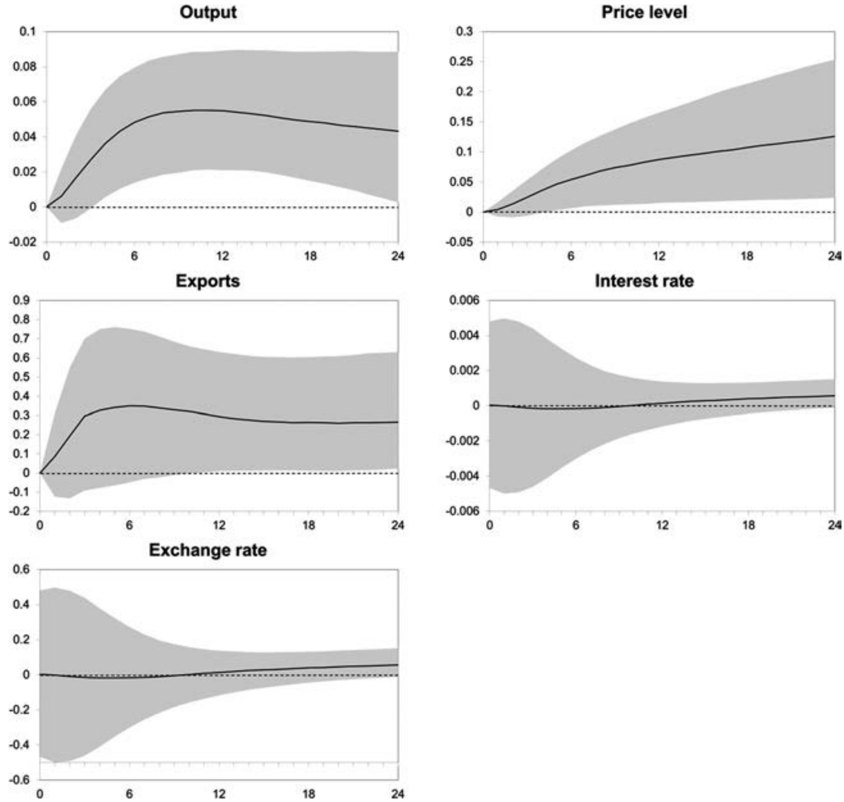


Notes: The figure shows the response of variables to an expansionary one-standard-deviation Eurosystem balance sheet shock. The shaded regions report pointwise 68 percent credibility intervals. The x-axis reports months, and the y-axis reports monthly growth rates in percent for all variables except for interest rates, which depict changes in percentage points. The results are derived from the baseline model (output and price level) as well as separate models including exports, the interest rate, and exchange rate.

is in line with the relative exchange rate stability of the lei vis-à-vis the euro since mid-2012.[12] The muted exchange rate response confirms the outcome of the event study by Falagiarda, McQuade, and

[12]From mid-2012 to end-2017 the average monthly fluctuation against the euro amounted to 0.5 percent, as compared with 1.0 percent from 2008 up to mid-2012.

Figure 11. Serbia: Response to a Balance Sheet Shock



Notes: The figure shows the response of variables to an expansionary one-standard-deviation Eurosystem balance sheet shock. The shaded regions report pointwise 68 percent credibility intervals. The x-axis reports months, and the y-axis reports monthly growth rates in percent for all variables except for interest rates, which depict changes in percentage points. The results are derived from the baseline model (output and price level) as well as separate models including exports, the interest rate, and exchange rate.

Tirpák (2015), while they also find a pronounced reaction of the short-term money market rate which is different from the results obtained here.

Serbia's output (see figure 11) increases, with a peak of 0.06 percent after eleven months, which is one of the strongest output responses compared with the other countries in the region. Also,

Table 4. Sign and Zero Restrictions for the Shock Identification of the Shadow Rate Model

$shadow_t^{EA}$	$CISS_t$	$spread_t^{EA}$	MRO_t	y_t^{EA}	p_t^{EA}	y_t^{SEE}	p_t^{SEE}
– 0–1	– 0–1	– 0–1	0	0	0	0	0

Note: 0 indicates that the immediate response is restricted, while + (–) indicates that only a positive (negative) reaction is permitted in the respective period.

the price response is very pronounced, with a peak increase of 0.13 percent at the end of the two-year horizon, which inter alia can be explained by the strong contribution of euro-area import prices to inflation pressures in the past (see, e.g., International Monetary Fund 2011). Serbia's exports seem to react positively to the shock, with a peak response of 0.35 percent after seven months. On the contrary, neither the interbank interest rate nor the exchange rate seems to be affected significantly by the shock.[13]

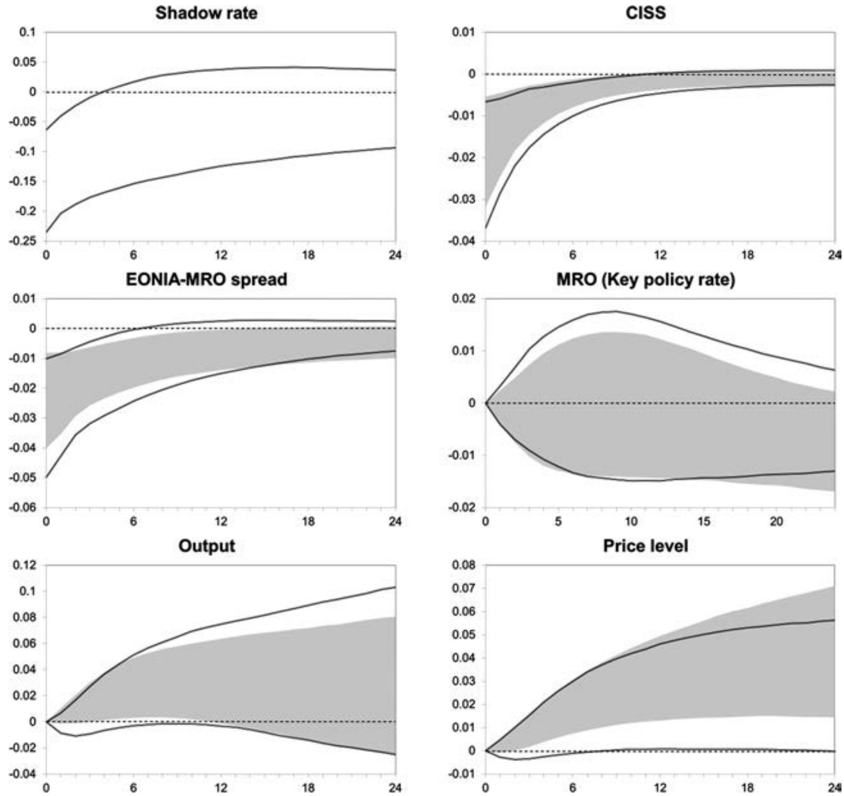
5. Robustness Testing

To test whether the results hold for different model specifications or variable choices, I perform the following robustness checks: As discussed in subsection 3.1, besides balance sheet assets, shadow rates can also be used as an indicator for non-standard monetary policy measures. To see whether the results are robust to an expansionary shadow rate shock (where I use the shadow rate developed by Wu and Xia 2016), I keep all other variables and the shock identification unchanged. The only difference is that an expansionary shock implies that the shadow rate *decreases*, which means that the sign restriction for the shadow rate is turned into negative (see table 4).

The results of a one-standard-deviation shadow rate shock for the euro area are depicted in figure 12. The double lines represent the credibility interval of the shadow rate shock, while the shaded area indicates the credibility interval of the baseline model. Compared

[13]This is despite the fact that, compared with Albania and Romania, Serbia's exchange rate fluctuated relatively strong vis-à-vis the euro in the sample period, with an average monthly fluctuation of 0.9 percent.

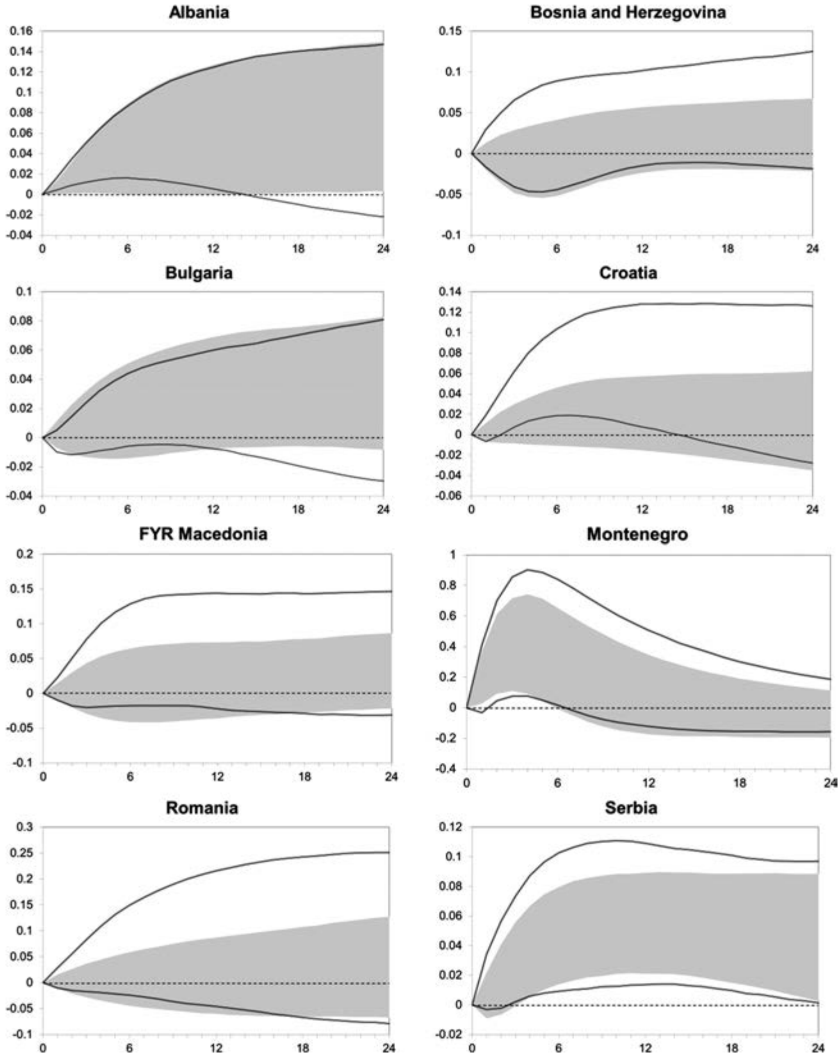
Figure 12. Euro Area: Response to a Shadow Rate Shock



Note: The shaded area represents the 68 percent credibility interval of the benchmark model, while the double lines indicate the credibility interval of the response to an expansionary one-standard-deviation shadow rate shock.

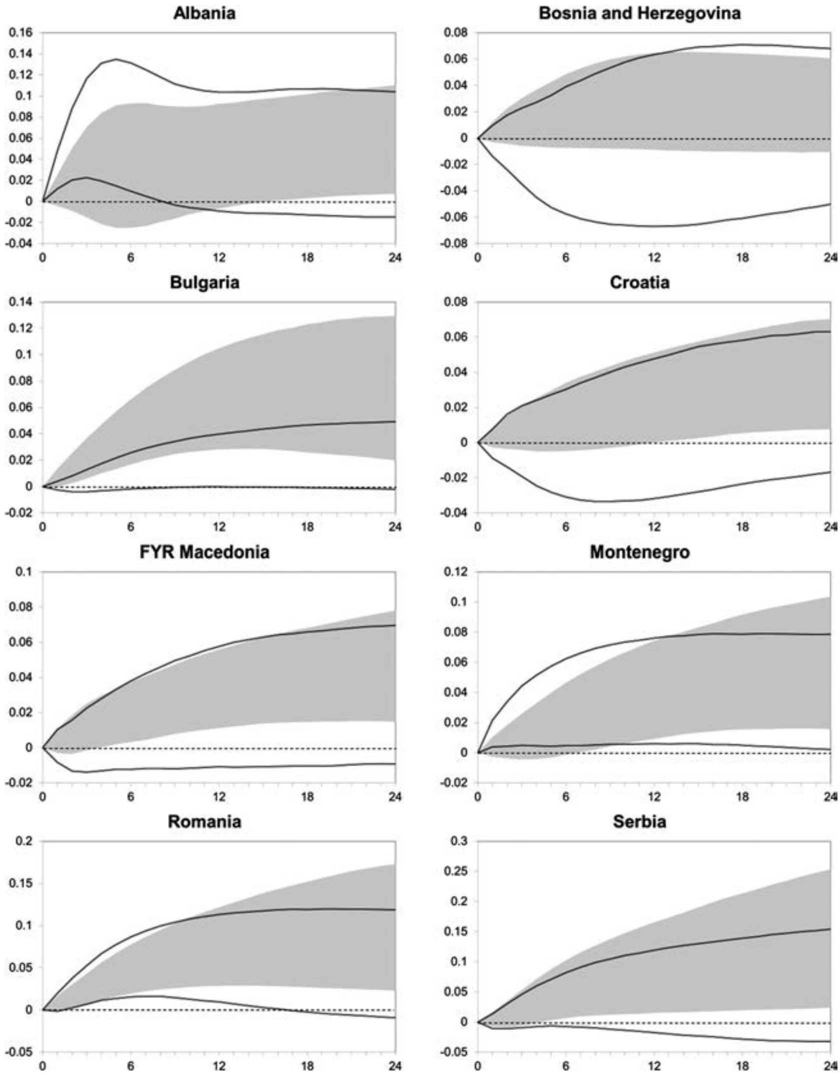
with a balance sheet shock (see figure 2), the shadow rate shock is less persistent. Regardless of the different time horizon, the responses of the financial variables (CISS, EONIA-MRO spread, and MRO) are very similar. The effects on output and prices are also in line with the baseline model, although the median price-level response is lower and the credibility intervals are larger when using the shadow rate as an indicator for non-standard monetary policy measures. The spillovers to output and price levels of the SEE countries are depicted in figures 13 and 14. The outcome is qualitatively in line with the

Figure 13. Response to a Shadow Rate Shock: Effect on Output



Note: The shaded area represents the 68 percent credibility interval of the benchmark model, while the double lines indicate the credibility interval of the response to an expansionary one-standard-deviation shadow rate shock. For Montenegro, industrial production is taken as a measure for output.

Figure 14. Response to a Shadow Rate Shock: Effect on Price Level



Note: The shaded area represents the 68 percent credibility interval of the benchmark model, while the double lines indicate the credibility interval of the response to an expansionary one-standard-deviation shadow rate shock.

baseline model. In some cases, the price-level response is less pronounced than in the baseline model, which can be explained by the more muted response of euro-area prices (see figure 12). Moreover, I reestimate all models by using only the position "Securities held for monetary policy purposes" (A070100) of the Eurosystem's balance sheet. As discussed in subsection 3.1, this balance sheet position reflects the most exogenous non-standard monetary policy measure and therefore serves as another robustness test for exogenous monetary policy shocks. Furthermore, I can also infer from the results whether spillovers of securities purchases are different from composite spillovers of all programs. Again, the results are qualitatively robust.

6. Conclusion

This paper is the first one to comprehensively assess the economic impact of the euro area's non-standard monetary policy measures on the countries of southeastern Europe (SEE). By employing bilateral structural BVAR models, I am able to identify macroeconomic spillovers as well as potential transmission channels for each country individually. Three questions are addressed in this paper: First, how have the ECB's non-standard monetary policy measures been affecting the SEE countries? Second, which channels are transmitting these shocks to SEE? Third, do different exchange rate regimes play a role in the SEE countries' responses to the shock?

The results show that the price level of all countries is positively affected by an expansionary non-standard monetary policy shock originating in the euro area, in line with the importance of euro-area imports in total imports. Compared with the euro-area response of the price level, the inflationary effect on SEE is larger in most SEE countries. With regard to the output response, the shock has an expansionary effect in approximately half of the countries, which is in some cases also more pronounced than the euro-area output response. These results are confirmed by robustness checks.

Regarding possible transmission channels, I find that spillovers seem to be mostly transmitted via the export channel. On the contrary, the interest rate channel exhibits a pronounced response only in a few countries, which might be driven by the relative illiquid interbank market in the SEE countries. Nevertheless, financial flows in the form of foreign direct or portfolio investments, which are not captured in the model, still might play a role.

With respect to the exchange rate regime, I find no influence of it on the price-level or output responses. This is in line with the absence of a distinct exchange rate response in the model output for the countries under a flexible regime, which can in turn be explained by the very stable exchange rate the respective currencies have exhibited vis-à-vis the euro in the past years.

The current work could be extended in various directions. A comparison between the spillovers of euro-area non-standard and euro-area conventional monetary policy measures could indicate whether these measures have different international effects. Moreover, future research might include additional variables to shed more light on potential transmission channels; especially the role of financial transmission could be further explored and other channels not covered here could be added (e.g., confidence channel). Finally, a comparison with spillovers from non-standard monetary policy measures undertaken by the central banks of other large advanced economies (notably the United States) would shed light on the relative importance of euro-area non-standard monetary policy, and be helpful for policymakers to design optimal policy responses to advanced economies' monetary policy measures and their (potential) reversal.

References

Arias, J. E., J. F. Rubio-Ramirez, and D. F. Waggoner. 2014. "Inference Based on SVAR Identified with Sign and Zero Restrictions: Theory and Applications." CEPR Discussion Paper No. 9796.

Babecká Kucharčuková, O., P. Claeys, and B. Vašíček. 2016. "Spillover of the ECB's Monetary Policy Outside the Euro Area: How Different is Conventional from Unconventional Policy?" *Journal of Policy Modeling* 38 (2): 199–225. doi: 10.1016/j.jpolmod.2016.02.002.

Benkovskis, K., A. Bessonovs, M. Feldkircher, and J. Wörz. 2011. "The Transmission of Euro Area Monetary Shocks to the Czech Republic, Poland and Hungary: Evidence from a FAVAR Model." *Focus on European Economic Integration* 2011 (3): 8–36.

Bluwstein, K., and F. Canova. 2016. "Beggar-thy-Neighbor? The International Effects of ECB Unconventional Monetary Policy

Measures." *International Journal of Central Banking* 12 (3, September): 69–120.

Boeckx, J., M. Dossche, and G. Peersman. 2017. "Effectiveness and Transmission of the ECB's Balance Sheet Policies." *International Journal of Central Banking* 13 (1, February): 297–333.

Burriel, P., and A. Galesi. 2018. "Uncovering the Heterogeneous Effects of ECB Unconventional Monetary Policies across Euro Area Countries." *European Economic Review* 101 (January): 210–29.

Canova, F. 2005. "The Transmission of US Shocks to Latin America." *Journal of Applied Econometrics* 20 (2): 229–51. doi: 10.1002/jae.837.

Cerutti, E., A. Ilyina, Y. Makarova, and C. Schmieder. 2010. "Bankers without Borders? Implications of Ring-Fencing for European Cross-Border Banks." IMF Working Paper No. 10/247.

Chen, Q., A. Filardo, D. He, and F. Zhu. 2012. "International Spillovers of Central Bank Balance Sheet Policies." *BIS Papers* 66 (September): 230–74.

———. 2015. "Financial Crisis, US Unconventional Monetary Policy and International Spillovers." BIS Working Paper No. 494.

Ciarlone, A., and A. Colabella. 2016. "Spillovers of the ECB's Non-standard Monetary Policy into CESEE Economies." Occasional Paper No. 351, Banca d'Italia.

Crespo Cuaresma, J., G. Doppelhofer, M. Feldkircher, and F. Huber. 2016. "US Monetary Policy in a Globalized World." Working Paper No. 205, Oesterreichische Nationalbank.

Cushman, D. O., and T. Zha. 1997. "Identifying Monetary Policy in a Small Open Economy under Flexible Exchange Rates." *Journal of Monetary Economics* 39 (3): 433–48. doi: 10.1016/S0304-3932(97)00029-9.

Dieppe, A., R. Legrand, and B. van Roye. 2016. "The BEAR Toolbox." ECB Working Paper No. 1934. doi: 10.2866/292952.

European Central Bank. 2016. "The International Role of the Euro." Technical Report.

Falagiarda, M., P. McQuade, and M. Tirpák. 2015. "Spillovers from the ECB's Non-standard Monetary Policies on Non-euro Area EU Countries: Evidence from an Event-Study Analysis." ECB Working Paper No. 1869. doi: 10.13140/RG.2.2.17327.74401.

Feldkircher, M. 2015. "A Global Macro Model for Emerging Europe." *Journal of Comparative Economics* 43 (3): 706–26. doi: 10.1016/j.jce.2014.09.002.

Fratzscher, M., M. Lo Duca, and R. Straub. 2016. "On the International Spillovers of US Quantitative Easing." *Economic Journal* 128 (608): 330–77. doi: 10.1111/ecoj.12435.

Friedman, M. 1953. "The Case for Flexible Exchange Rates." In *Essays in Positive Economics*. University of Chicago Press.

Gambacorta, L., B. Hofmann, and G. Peersman. 2014. "The Effectiveness of Unconventional Monetary Policy at the Zero Lower Bound: A Cross-Country Analysis." *Journal of Money, Credit and Banking* 46 (4): 615–42. doi: 10.1111/jmcb.12119.

Georgiadis, G. 2016. "Determinants of Global Spillovers from US Monetary Policy." *Journal of International Money and Finance* 67 (October): 41–61. doi: 10.1016/j.jimon_n.2015.06.010.

Georgiadis, G., and J. Gräb. 2015. "Global Financial Market Impact of the Announcement of the ECB's Extended Asset Purchase Programme." Working Paper No. 232, Globalization and Monetary Policy Institute, Federal Reserve Bank of Dallas.

Hájek, J., and R. Horváth. 2016. "The Spillover Effect of Euro Area on Central and Southeastern European Economies: A Global VAR Approach." *Open Economies Review* 27 (2): 359–85. doi: 10.1007/s11079-015-9378-4.

Halova, K., and R. Horváth. 2015. "International Spillovers of ECB's Unconventional Monetary Policy: The Effect on Central and Eastern Europe." IOS Working Paper No. 351, Institut für Ost- und Südosteuropaforschung.

Holló, D., M. Kremer, and M. Lo Duca. 2012. "CISS — A Composite Indicator of Systemic Stress in the Financial System." ECB Working Paper No. 1426. doi: 10.2139/ssrn.1611717.

International Monetary Fund. 2011. "Republic of Serbia: Seventh Review and Inflation Consultation under the Stand-By Arrangement." IMF Country Report No. 11/95.

Jarociński, M. 2010. "Responses to Monetary Policy Shocks in the East and the West of Europe: A Comparison." *Journal of Applied Econometrics* 25 (5): 833–68. doi: 10.1002/jae.1082.

Jiménez-Rodriguez, R., A. Morales-Zumaquero, and B. Ègert. 2010. "The Effect of Foreign Shocks in Central and Eastern Europe." *Journal of Policy Modeling* 32 (4): 461–77.

Krippner, L. 2015. "A Comment on Wu and Xia (2015), and the Case for Two-Factor Shadow Short Rates." Working Paper No. 48/2015, Centre for Applied Macroeconomic Analysis.

Lombardi, M., and F. Zhu. 2014. "A Shadow Policy Rate to Calibrate US Monetary Policy at the Zero Lower Bound." BIS Working Paper No. 452.

Maćkowiak, B. 2006. "How Much of the Macroeconomic Variation in Eastern Europe is Attributable to External Shocks?" *Comparative Economic Studies* 48 (3): 523–44. doi: 10.1057/palgrave.ces.8100143.

———. 2007. "External Shocks, U.S. Monetary Policy and Macroeconomic Fluctuations in Emerging Markets." *Journal of Monetary Economics* 54 (8): 2512–20. doi: 10.1016/j.jmoneco.2007.06.021.

Meade, J. 1951. *The Theory of International Economic Policy*. 2 vols. Oxford University Press.

Minea, A., and C. Rault. 2011. "External Monetary Shocks and Monetary Integration: Evidence from the Bulgarian Currency Board." *Economic Modelling* 28 (5): 2271–81. doi: 10.1016/j.econmod.2011.05.008.

Miniane, J., and J. H. Rogers. 2007. "Capital Controls and the International Transmission of U.S. Money Shocks." *Journal of Money, Credit and Banking* 39 (5): 1003–35. doi: 10.1111/j.1538-4616.2007.00056.x.

Peersman, G. 2011. "Macroeconomic Effects of Unconventional Monetary Policy in the Euro Area." ECB Working Paper No. 1397.

Petrevski, G., P. Exterkate, D. Tevdovski, and J. Bogoev. 2015. "The Transmission of Foreign Shocks to South Eastern European Economies: A Bayesian VAR Approach." *Economic Systems* 39 (4): 632–43.

Potjagailo, G. 2017. "Spillover Effects from Euro Area Monetary Policy across the EU: A Factor-Augmented VAR Approach." *Journal of International Money and Finance* 72 (April): 127–47. doi: 10.1016/j.jimon_n.2017.01.003.

Sahay, R., V. Arora, T. Arvanitis, H. Faruqee, P. N'Diaye, and T. Mancini-Griffoli. 2014. "Emerging Market Volatility: Lessons from the Taper Tantrum." Staff Discussion Note No. 14/09, International Monetary Fund.

Wu, J. C., and F. D. Xia. 2016. "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound." *Journal of Money, Credit and Banking* 48 (2–3): 253–91. doi: 10.1111/jmcb.12300.

Effects of Changing Monetary and Regulatory Policy on Money Markets*

Elizabeth Klee, Zeynep Senyuz, and Emre Yoldas
Federal Reserve Board

The global financial crisis and the resulting policy response led to substantial changes in U.S. dollar funding markets, which are crucial for the functioning of the financial system and the transmission of monetary policy in the United States. We develop and test hypotheses on the effects of changing monetary and regulatory policy on key funding rates. We show that the federal funds rate continued to provide an anchor for unsecured rates, albeit weaker, while its transmission to the secured repo rate is hampered in the post-crisis period. The Federal Reserve's reverse repurchase facility led to stronger co-movement and reduced volatility of money market rates. The new regulations and the superabundant reserves environment affected rate dynamics on calendar days primarily through increased balance sheet costs.

JEL Codes: C32, E43, E52, G21, G28.

1. Introduction

During the global financial crisis of 2007 to 2009, and in the years following, the Federal Reserve (Fed) injected massive amounts of

*We would like to thank James Clouse, Jane Ihrig, Fulvio Pegoraro, and participants at the 2017 European Central Bank Money Markets Workshop, 2016 International Finance and Banking Society Barcelona Conference, 2016 Western Economic Association International Conference, and UC Riverside Economics Department seminar for helpful comments. We thank Richard Sambasivam for his assistance with the data set and Gurubala Kotta for her editorial suggestions. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System, or other members of its staff. Author e-mails: elizabeth.c.klee@frb.gov; zeynep.senyuz@frb.gov; emre.yoldas@frb.gov.

liquidity into the financial system, kept its target policy rate near zero, and introduced new tools to conduct monetary policy. Meanwhile, a series of regulatory reforms that prompted financial institutions to reevaluate their risk-management practices were announced and implemented. These changes to monetary policy implementation and the regulatory environment marked a new era for the financial system. In this paper we focus on the impact of these developments on money markets.

Understanding the effects of changes to the Fed's monetary policy framework and financial regulations is important for a number of reasons. To start, the Fed uses money markets to influence broader financial conditions. Specifically, the first step of monetary policy transmission is for the Fed's policy rate (the federal funds rate) to influence other overnight interest rates. For effective policy implementation, other overnight rates should move closely with the federal funds rate, that is, after controlling for risk factors and market frictions, rate differentials should be arbitraged away. Impeded monetary policy transmission could pose challenges to the central bank for controlling interest rates. Therefore, it is important to identify how and to what extent pass-through from the federal funds rate to other money market interest rates has been affected since the crisis.

In addition, some adjustments to the regulatory framework that were aimed at money markets also created new incentives for market participants. Reliance of dealers and other financial intermediaries on short-term borrowing in money markets may have contributed in part to the financial crisis. Consequently, many regulations focused on reducing risks to institutions operating in money markets. Understanding the intended or unintended consequences of these actions helps policymakers evaluate the costs and benefits of these policies.

Finally, money markets represent a significant share of financial intermediation. Therefore, quantifying the effects of changes in monetary policy implementation and regulation on these markets is important. To provide context, at the end of the first quarter of 2018, there were roughly \$4.5 trillion in money market instruments outstanding, with federal funds and repurchase agreements (repos) alone representing \$3.5 trillion. As a point of comparison, there were about $5.4 trillion outstanding in non-financial corporate bonds in the same period.[1]

Against this backdrop, we document significant changes to the Fed's monetary policy implementation framework and to the regulatory environment over the course of the financial crisis and the subsequent effective lower bound (ELB) period. We then identify their effects on dynamics of key overnight funding rates. To do so, we estimate dynamic multivariate models for money market rates over two sample periods: the pre-crisis period that runs from January 2001 to 2007 and serves as a benchmark, and the ELB period from December 2008 to August 2015, during which the aforementioned policy changes took place. Our pre-crisis model incorporates the long-run relationship of the federal funds rate with the other overnight rates in a multivariate framework. We assume time variation in the volatilities and correlations of funding rates to allow for potentially different dynamics around financial reporting days, when some institutions contract their balance sheets.

Our results suggest that despite important changes in the market structure, the federal funds rate continued to provide an anchor for unsecured overnight rates. At the same time, the co-movement of rates weakened significantly. In particular, transmission of the federal funds rate to the repo rate was hampered. We also illustrate some differences in behavior for unsecured versus secured money market rates. Specifically, we show that the new regulations substantially altered unsecured rate dynamics on financial reporting days by increasing balance sheet costs of financial intermediaries. Rates that represent unsecured wholesale funding costs for banks became significantly lower and more volatile on quarter-ends. By contrast, for secured rates, quarter-end effects weakened in the repo market, reflecting lower dealer leverage in response to new regulations and reduced net repo financing. Separately, day-of-week effects on the federal funds rate mostly disappeared due to the abundance of reserve balances post-crisis.

In addition to exploring differences between the pre-crisis and ELB periods, we also take a close look at structural changes during the post-crisis period. We use September 2013 as a natural

[1]Refer to Federal Reserve Z.1 release, Financial Accounts, tables 209 and 213.

breakpoint in this sample, as two major developments took place around that time. First, the Fed expanded its monetary policy toolkit with the introduction of the overnight reverse repurchase (ON RRP) facility. Second, a number of Basel III regulatory changes were announced around that time. We document that the ON RRP facility strengthened the link between the repo rates and unsecured rates, and also contributed to better monetary policy transmission. Moreover, volatility of all rates dampened, with an especially notable decline in the repo market. We also find that the tendency of foreign banks to reduce their overnight borrowing on financial-reporting-related days, combined with the search by cash lenders for alternative investment opportunities, exacerbated month-end and quarter-end effects on the federal funds rate and Eurodollar rates. The availability of the ON RRP as a viable investment option on financial reporting days, when alternatives are limited, reduced the potential for sharp drops in the repo rate.

In related literature on money market dynamics, Afonso, Kovner, and Schoar (2011) analyze activity in the federal funds market during the global financial crisis, while Gorton and Metrick (2012) and Copeland, Martin, and Walker (2014) focus on the role of repo markets during the crisis in the context of runs. Yoldas and Senyuz (2015) model the behavior of term money market rates and quantify financial stress. Although the literature on monetary policy transmission to the economy is vast, there is relatively limited research on how the target rate is transmitted to other overnight interest rates. Bech, Klee, and Stebunovs (2014) find evidence of deterioration of the pass-through from the federal funds rate to the repo rate during the financial crisis that seemed to persist, while Kroeger and Sarkar (2016) suggest that this pass-through improved with the ON RRP facility.

Another strand of literature related to our work documents the effects of calendar days on money market rates. Spindt and Hoffmeister (1988), Griffiths and Winters (1995), Hamilton (1996), Carpenter and Demiralp (2006), and Judson and Klee (2010) show that the federal funds rate exhibits calendar-day effects associated with the maintenance period as well as quarter-ends. More recently, Munyan (2015) and Anbil and Senyuz (2016) document the effects of window-dressing activity on financial reporting dates in the repo market.

The rest of the paper is organized as follows. In the next section, we provide background information on the mechanics of money markets, review changes in the monetary policy implementation framework and relevant regulations, and establish hypotheses on their effects. We describe the data set in section 3 and lay out the methodological framework in section 4. We present and discuss the estimation results in section 5. We conclude in section 6.

2. Money Markets and the Changing Landscape

The response of the Fed to the global financial crisis of 2007–09 significantly altered the money market landscape in the United States. With successive rate cuts, the target federal funds rate was reduced from 5.25 percent in August 2007 to its effective lower bound (ELB) of 0 to 0.25 percent in December 2008. The federal funds rate, as well as other overnight rates, remained at the ELB for the next seven years. Throughout the crisis and its aftermath, the Fed used a variety of new facilities to provide liquidity to the financial system as well as unconventional tools, such as large-scale asset purchases. As a result, reserves in the banking system have reached unprecedented levels. Marking a significant shift in its policy framework, the Fed started paying interest on reserves (IOR) in October 2008 to control rates in an environment of superabundant reserves.

Elevated reserves and the new monetary policy tools affected trading dynamics of money market participants. In the federal funds market, government-sponsored enterprises (GSEs), which are not eligible to earn IOR, became the primary lenders, while large and foreign banks started borrowing funds at rates below IOR for arbitrage purposes. Mainly because of this fragmented market structure, the IOR could not set a lower bound on the federal funds rate, leading the Fed to introduce a supplementary tool, the ON RRP facility, in September 2013 to enhance monetary control.

The changing regulatory environment also created new incentives for market participants amid a substantial decline in the leverage of securities dealers. Among the new regulations, the change in the assessment base for the Federal Deposit Insurance Corporation (FDIC) deposit insurance and the Basel III leverage ratio requirements are of particular importance. The former made wholesale funding more costly for U.S.-chartered banks relative to that of U.S.

branches and agencies of foreign banks, incentivizing domestic banks to reduce their money market borrowing. The latter incentivized foreign banks to dynamically deleverage through money market activity given regional differences in implementation of the leverage ratio. Meanwhile, both leverage levels and net repo liabilities of the broker-dealer sector decreased notably, creating an important contrast to the pre-crisis period, during which such institutions largely operated outside of the regulated banking system.

The remainder of this section describes the pre- and post-crisis monetary policy and regulatory frameworks in more detail to give perspective on the empirical analysis that follows. We cover different segments of the money market in sections 2.1 and 2.2, provide background on the Fed's monetary policy implementation framework in section 2.3, and discuss regulations in section 2.4.

2.1 Bank Reserves and Activity in the Federal Funds Market

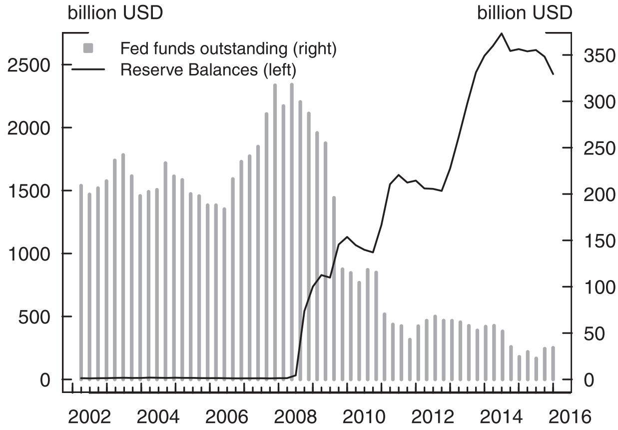
Banks are required to maintain a minimum level of reserves at the Federal Reserve Banks in their Districts.[2] Historically, banks avoided holding excess reserves, as such balances did not earn any interest. Indeed, total reserve balances in the banking system averaged about \$10 billion in 2007, with only about \$2 billion in excess reserve balances, while total bank assets were close to \$10 trillion over the same period. As can be seen in figure 1, reserves in the system increased to more than \$800 billion at the end of 2008, as the Fed provided liquidity during the financial crisis through several facilities.[3] Following subsequent rounds of asset purchases from 2009 to 2014, total reserve balances averaged nearly \$2.5 trillion in 2016.[4]

[2]We will be referring to depository institutions with reserve accounts simply as banks. See Regulation D Reserve Requirements for a full list of financial institutions in this category, available at https://www.federalreserve.gov/boarddocs/supmanual/cch/int_depos.pdf.

[3]See https://www.federalreserve.gov/monetarypolicy/bst_crisisresponse.htm for details on the Fed's crisis response, and https://www.federalreserve.gov/monetarypolicy/expiredtools.htm for a list of expired liquidity provision facilities.

[4]Between November 2008 and October 2014, the Fed purchased nearly \$1.7 trillion in Treasury securities and about \$2 trillion in agency mortgage-backed securities, as well as \$170 billion in agency debt securities in order to put downward pressure on longer-term interest rates. See Krishnamurthy and Vissing-Jorgensen (2011) and D'Amico et al. (2012) for a discussion of the economic rationale and effects of large-scale asset purchases.

Figure 1. Reserves and Federal Funds



Note: Data are quarterly and obtained from Call Reports.

The unprecedented increase in reserve balances changed the landscape for the federal funds market. Federal funds are unsecured loans of reserve balances between banks and other eligible institutions, mainly GSEs. Federal funds transactions are typically conducted for an overnight term and are carried out either directly between the institutions or through third-party brokers. Historically, transactions in the federal funds market facilitated the redistribution of reserve balances, whereby banks with reserve balances in excess of the required levels lent to banks in need of reserves. The surge in reserve balances led to a substantial decline in banks' need for short-term borrowing to cover idiosyncratic funding shortfalls. To ensure monetary control and promote efficiency in the banking system, the Fed introduced the IOR as a new monetary policy tool in October 2008. As a result, incentives for banks to lend federal funds at rates below the IOR rate were largely eliminated.

As shown in figure 1, the end-of-quarter outstanding amount of federal funds borrowed by banks declined to roughly one-fourth of the level observed prior to the global financial crisis by 2011, and it has remained low since then. Moreover, volume in the overnight federal funds market declined from an average of $200 billion per day in 2007 to $60 billion per day at the end of 2012, according to estimates provided by Afonso, Entz, and LeSueur (2013b). Banks that used to lend more than half of the federal funds before the crisis accounted for only a fraction of the lending activity after 2008. GSEs, specifically the Federal Home Loan Banks, which are not eligible to earn IOR, have been the main lenders in the post-crisis period.[5] On the borrowing side, Afonso, Entz, and LeSueur (2013a) show that mostly banks under the umbrella of bank holding companies and foreign banking organizations have been purchasing federal funds from GSEs for arbitrage purposes.[6] These institutions borrow federal funds at rates below the IOR and place the cash in their reserve accounts to earn the spread between the IOR and the federal funds rate. These transactions have been relatively more profitable for foreign banks, as they are not subject to assessment by the FDIC.[7]

The changing landscape in the federal funds market raises the important question of whether pass-through from the federal funds rate to other overnight rates has been affected over time. In addition, superabundant reserves may have implications for cash flows in the market associated with days of reserve maintenance period.[8] In the pre-crisis era, activity in the federal funds market was in part driven by maintenance-period dynamics, as shown by Spindt and Hoffmeister (1988), Griffiths and Winters (1995), and Hamilton (1996), among others. However, superabundant reserves, combined with the finding by Ennis and Wolman (2015) that reserves in the system have been fairly widely distributed across banks since mid-2009, suggest that calendar effects associated with reserve maintenance have likely dissipated in the post-crisis era.

[5]See Ashcraft, Bech, and Frame (2010) for a detailed description of the Federal Home Loan Bank System and its role as a liquidity provider to banks.

[6]In the current context, foreign banking organizations are U.S. branches and agencies of foreign banks.

[7]In 2011, the FDIC changed the assessment base for its deposit insurance scheme from domestic deposits to total assets minus equity, making larger balances more costly for domestic banks regardless of funding source. See Kreicher, McCauley, and McGuire (2013) for a detailed discussion.

[8]See the Reserve Maintenance Manual for reporting requirements as well as calculation and maintenance of reserve balances, available at http://www.federalreserve.gov/monetarypolicy/files/reserve-maintenance-manual.pdf.

2.2 Other Money Market Segments

The Fed's monetary policy implementation framework relies on targeting the federal funds rate and cross-market linkages across money markets. Specifically, overlapping participants in various money market segments and arbitrage activity ensure strong co-movement that the Fed depends on for effective monetary policy transmission.

The Eurodollar market is another segment for unsecured funding that is broader than the federal funds market. Eurodollars are U.S. dollar-denominated deposits held in a bank or a bank branch located outside of the United States. U.S. banks and foreign banking organizations cannot directly borrow in the Eurodollar market but can take Eurodollar deposits, mainly through their Caribbean branches, and transfer them onshore to fund U.S. operations. Eurodollar deposits that remain outside the United States are not covered by FDIC deposit insurance, while those that are transferred to an insured U.S. affiliate are included in the deposit insurance assessment base. Because of their unsecured nature and regulatory treatment, Eurodollar deposits constitute a close substitute to federal funds. However, the Eurodollar market has a more diverse set of participants than the federal funds market, as participants do not have to have an account at the Fed. Cipriani and Gouny (2015) estimate that the average volume in the brokered Eurodollar market is three to four times larger than that in the brokered federal funds market. However, there has been a substantial drop in the Eurodollar volume following the money fund reform compliance date in October 2016, as prime funds pulled back from lending in this market.

The major segment of the money market for secured funding is the repo market. A repo is effectively a collateralized loan in which the lender of the cash receives the security as collateral, and the borrower pays the lender interest on the loan. Sale of securities takes place under an agreement to repurchase them at a specified price on a later date.[9] The repo market can broadly be divided into two

[9]Fed transactions in the repo market are defined from the point of view of the market participants, that is, a collateralized transaction in which the Fed borrows cash is called a reverse repo.

parts: the bilateral market where the two parties interact directly, and the triparty market where clearing/brokerage services of a third party are involved. Total volume of the Treasury repo market is well above $2 trillion. Copeland et al. (2014) and Baklanova et al. (2016) provide estimates of volumes in different repo market segments. Cash borrowers (or securities lenders) in the repo market include banks and securities dealers, while money market mutual funds (or money funds) and GSEs are among the primary lenders of cash (or borrowers of securities).

The final segment of the money market we consider is the commercial paper market, in which large corporations issue debt for a fixed maturity. Many companies issue commercial paper when they need to raise short-term cash, as it is a lower-cost alternative to bank loans. Although commercial paper is unsecured, it is considered a very safe investment, as typically only creditworthy companies with high credit ratings issue such securities. Commercial paper is especially attractive for institutional investors, as they are liquid and have a low risk of default.

2.3 Monetary Policy Implementation Framework

Historically, adjustment of the level of reserve balances in the banking system to move the effective federal funds rate toward the target level set by the Federal Open Market Committee (FOMC) was the central pillar of monetary policy implementation. Given scarce reserve balances in the system, the Fed would affect the federal funds market rate by announcing a target level and managing the amount of reserves available to the banking system through open market operations (OMOs). These operations would influence the rate in the federal funds market, where banks experiencing shortfalls could borrow from banks with excess reserves. Given the low volume of reserves at the Fed, around $10 billion, even small OMOs could significantly affect the market rate. Changes in the federal funds rate would then be transmitted to other short-term interest rates, to longer-term interest rates, and eventually to inflation and economic activity. This framework worked seamlessly while the Fed was operating with a balance sheet of less than $1 trillion before the crisis.

The global financial crisis precipitated changes in the operational framework of the Fed.[10] In an environment with superabundant reserves, the conventional approach based on changing the quantity of reserves via OMOs would not work. As a result, the Fed extended its monetary policy toolkit. In the fall of 2008, the Fed started paying interest on banks' reserve balances, which became the primary tool of its new monetary policy implementation framework in controlling short-term interest rates.

Although adjusting the IOR rate is an effective way to move market interest rates in an environment of superabundant reserves, federal funds have generally traded below this rate, mainly due to the fact that only banks can earn IOR. GSEs, the other major group of participants in the federal funds market, still have an incentive to lend at rates below the IOR rate, as they do not receive interest on their reserve accounts. Moreover, FDIC fees and other balance sheet constraints, such as capital and liquidity regulations, limit arbitrage activity by banks that would push the market rate toward the IOR rate.

In order to enhance monetary control and put an effective floor under short-term interest rates, the Fed introduced the ON RRP facility as a supplementary tool for its implementation of monetary policy. ON RRPs are offered to a broader set of financial institutions, including money funds that do not have access to the federal funds market. In September 2014, the FOMC issued a statement summarizing the new operating strategy, and in December 2015, it successfully lifted the federal funds rate from its near-zero range in this framework.[11]

The primary tool of the new operating framework, IOR, has important implications for the transmission of monetary policy from federal funds to the repo market. In the pre-crisis era, the active presence of large banks in both the federal funds and repo markets was crucial to the co-movement of these two rates. The unsecured nature of the federal funds transactions has typically resulted in a small and

[10]See Ihrig, Weinbach, and Meade (2015) for an in-depth discussion of the evolution of the Fed's monetary policy implementation framework through the financial crisis and its aftermath.

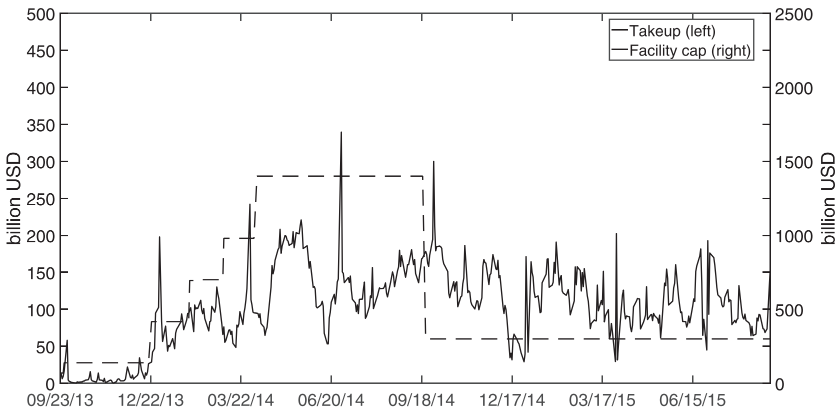
[11]See https://www.federalreserve.gov/monetarypolicy/policy-normalization.htm for further details on policy normalization. Anderson et al. (2016) provide an overview of money market developments after the liftoff.

positive spread between the federal funds rate and the rate on repo transactions where the underlying collateral is a U.S. Treasury or agency security. However, market-determined, or effective, federal funds rate staying below the repo rates became a frequent phenomenon amid superabundant reserves and the ELB on the funds rate. The negative spread reflects reduced scope for arbitrage activity due to IOR, aside from the dramatic reduction in banks' needs for short-term borrowing, as discussed previously. Specifically, when the repo rates were greater than the federal funds rate in the past, banks could borrow in the federal funds market and place the cash in the repo market, creating downward pressure on the repo rates and pushing the effective federal funds rate up. However, in the presence of the IOR, the incentive for banks to engage in arbitrage activity across the federal funds and repo markets exists only when the repo rates are above the IOR. Although GSEs may also engage in this type of arbitrage, frictions—such as internal restrictions or intraday timing considerations—likely limit such activity. As a result, we expect a weaker link between the effective federal funds rate and the repo rates in the ELB sample, on net.

The supplementary monetary policy tool of the new framework, the ON RRP facility, has also been affecting overnight funding dynamics since its inception in September 2013. The Fed has been conducting operations on a daily basis at a pre-announced offering rate. Through this facility, the Fed borrows cash from eligible counterparties in exchange for Treasury securities from its open market portfolio. These operations provide an investment vehicle for money market participants who usually compare the facility's offering rate with rates in the market and determine whether to bid in the operation.

The ON RRP operations are, in essence, similar to the temporary OMOs in the form of reverse repos conducted by the Fed prior to the crisis, but there are also important differences. Participation in the ON RRP operations is open to a wide range of entities, including money funds, banks, and GSEs, in addition to primary dealers of the Fed. Frost et al. (2015) show that money funds have been the dominant cash lenders in ON RRP operations. By expanding the set of alternative investments available to money funds and GSEs, the ON RRP was intended to strengthen rate control. The second important difference between the ON RRP and conventional

Figure 2. ON RRP Operations



Note: Data are daily and obtained from the Federal Reserve Bank of New York, available at https://apps.newyorkfed.org/markets/autorates/temp.

temporary OMOs is that the latter was conducted to move the effective federal funds rate close to the FOMC's target, while the former is intended to set a floor for the overnight rates. The mechanism is similar to that of IOR for banks in the federal funds market; ON RRP counterparties do not have an incentive to invest in alternative sources unless they are offered the ON RRP rate or higher. Indeed, Potter (2015) shows that the ON RRP has established a *soft* floor, as the FOMC intended—that is, although some trades likely occur below the ON RRP rate, volume-weighted average overnight funding rates have mostly been above the offering rate. A general reduction in the volatility of overnight rates is another expected effect of the soft floor set by the ON RRP. Such effects are especially important on financial reporting days when borrowers contract the size of their balance sheets, leaving cash lenders looking for alternative safe investment options.

Take-up at the ON RRP facility trended up for about a year following its inception in September 2013, as can be seen from figure 2. One year later, the FOMC reduced the overall limit on the facility substantially (from $1.4 trillion to $300 billion) and introduced an auction process to allocate reverse repos in the event that the overall limit is binding. This change led to a sharp drop in money market rates on that quarter-end, as it left cash lenders scrambling for alternative investments. In October 2014, term RRP operations spanning year-end were announced. These operations were conducted over quarter-ends until December 2015 and helped address downward pressure on rates, suggesting that perceived investment capacity is an important factor in determining the effectiveness of reverse repos in supporting rates.

At the time of the rate hike in December 2015, the aggregate cap on ON RRP operations was temporarily suspended. Currently, ON RRP operations are limited only by the value of Treasury securities in the Fed's open market portfolio that are available for these operations, which stand around $2 trillion.

2.4 New Banking Regulations and Dealer Leverage

The announcement and implementation of Basel III capital and liquidity reforms had a significant effect on the post-crisis financial landscape. In terms of their effects on money markets, the liquidity coverage ratio (LCR) and the leverage ratio are of particular interest among the Basel III reforms.

The LCR rule, which was first proposed in the United States in October 2013, requires banks to hold high-quality liquid assets (HQLAs) to meet cash outflows under a thirty-day stress scenario. Therefore, it has potential implications for bank activity in overnight money markets, as many assets and liabilities closely tied to these markets are under the jurisdiction of the LCR.

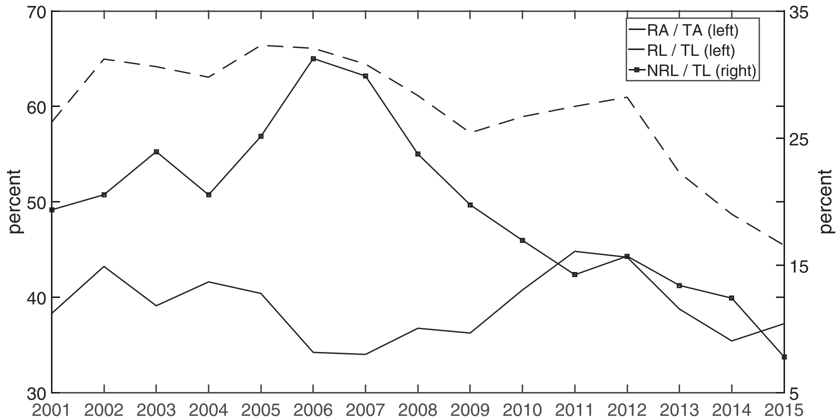
In the LCR calculation, cash inflows and outflows over a thirty-day stress period are aggregated and netted with specific rates for different assets and liabilities. In some cases, lending in the federal funds market may decrease the LCR of the lending bank. Although lending federal funds reduces the LCR numerator because reserves are counted as HQLAs, cash inflow assumptions for financial institutions typically imply limited impact of such activity on LCR. Similarly, lending in the repo market in which the underlying collateral is in the HQLA category has no effect on a bank's LCR. On the borrowing side, funding non-HQLAs through unsecured interbank borrowing or repos deteriorates LCR, and thus incentivizes banks to reduce their reliance on such financing. By the time the initial LCR announcement was made, banks had already reduced their reliance on wholesale funding substantially (see, for example, Choi and Choi 2016). Finally, IOR arbitrage trades increase a bank's LCR, as the borrowed cash is parked in the arbitrageur bank's reserve account—which is treated as HQLAs with no haircuts—and the cash outflow assumption associated with borrowing results in a less-than-proportional increase in the denominator. All told, we do not expect the effect of the LCR to be material for overnight money market dynamics.

Another notable aspect of Basel III for money market activity has been the leverage ratio requirement, which requires banks to hold tier 1 equity equivalent to at least 3 percent of their leverage exposure calculated using their on- and off-balance-sheet assets, including reserves. The Supplementary Leverage Ratio, the regulation that implements the leverage ratio provisions in the United States, bases the relevant calculations on averages of *daily* values for on-balance-sheet items. In contrast, for most foreign banks, disclosures are based on month- or quarter-end levels. This regional implementation difference incentivized foreign banks to contract their balance sheets on financial reporting days and expand on other days. The leverage ratio requirements were announced in mid-2013, and banks started disclosing their leverage ratios to the public in January 2015, including three quarters of historical data. Becoming compliant before the beginning of public disclosures was an important motivation for banks to adjust their balance sheets. Therefore, we expect stronger reporting-day effects on rates in the ELB sample following the introduction of the leverage ratio requirements.

Declining leverage of securities broker-dealers has been another important feature of the post-crisis landscape (Adrian et al. 2013). Dealers were not subject to leverage limits prior to the crisis, as they were outside the regulated banking system. However, four out of the five major standalone investment banks with dealer arms have been integrated into bank holding companies via either acquisitions or conversions. This change has been among the main drivers of lower dealer leverage along with generally increased risk aversion in the aftermath of the crisis.

Dealers dynamically adjust their balance sheets mainly through short-term borrowing in the form of repos, as discussed in Adrian and Shin (2010). Along with overall leverage, repo activity of

Figure 3. Repo Financing Activity by Securities Brokers and Dealers



Notes: Data are annual and obtained from the Financial Accounts of the U.S. statistical release (Z.1) of the Federal Reserve Board, available at http://www.federalreserve.gov/releases/z1/. TA (TL) denotes total assets (liabilities), RA (RL) denotes repo assets (liabilities), and NR is net repo financing, that is, RL − RA.

dealers also declined relative to the pre-crisis norms. As can be seen from figure 3, although repo lending by dealers has been relatively stable since 2001, their repo borrowing has been lower since 2007. The change in *net* repo financing has been substantial: the ratio of net repo liabilities to total liabilities for dealers has been steadily decreasing since its peak in 2007, and reached about 8 percent in 2015, almost one-fourth of its level in 2007. Meanwhile, the ON RRP facility limited downward pressure on the repo rates around financial reporting days by setting a soft floor. Therefore, we expect weaker quarter-end effects on repo rates in the ELB sample. Table 1 summarizes all the aforementioned changes in the monetary policy and the regulatory environment, as well as their anticipated effects on overnight money markets.

3. Data

Our sample covers two main periods: the pre-crisis period that spans from January 2, 2001, to July 31, 2007; and the ELB period that runs

Table 1. Changes in Monetary and Regulatory Policy and Implications

Changes in Monetary and Regulatory Policy	**Anticipated Effects on Overnight Money Markets**
Superabundant Reserves and IOR	
Lower trading volumes in the federal funds market	(i) Weaker co-movement of EFFR with other rates (ii) Increased EFFR volatility
Reduced scope for EFFR-RPR arbitrage trades by banks	Weaker EFFR-RPR co-movement
Widespread distribution of reserves	MP effects diminish in the aggregate
ON RRP	
Inclusion of money funds and GSEs among counterparties	(i) Stronger co-movement of overnight interest rates (ii) Lower interest rate volatility (iii) Weaker financial reporting effects on RPR
New Regulations and Lower Dealer Leverage	
LCR	IOR arbitrage trades more attractive, but limited effect due to other regulatory constraints
FDIC assessment base change Leverage ratio	Stronger financial-reporting-day effects on unsecured rates and their volatility
Diminishing leverage and repo financing by dealers	Weaker financial-reporting-day effects on RPR

from December 17, 2008, to August 28, 2015.[12] The former is associated with the conventional monetary policy operating framework and serves as a benchmark, while the latter is a period during which

[12]We exclude the period from mid-2007 to late 2008 from our analysis, as this period is associated with unprecedented movements in the rates driven by the financial crisis.

overnight money markets were subject to the significant changes discussed above.

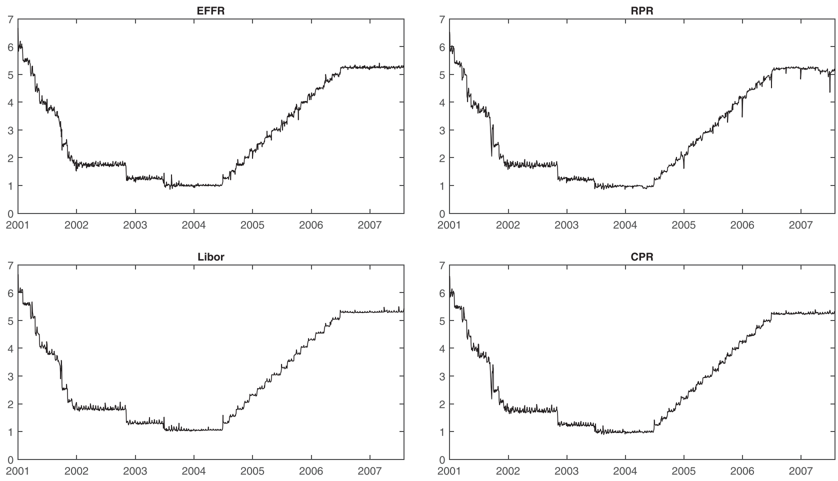
Our data set consists of four overnight money market interest rates. The first one is the effective federal funds rate (EFFR), which is calculated as a volume-weighted average of rates on brokered federal funds trades and published by the Federal Reserve Bank of New York (FRBNY). The second one is the Eurodollar rate (EDR), which represents the cost of alternative unsecured funding for large banks. We use the EDR data that the FRBNY started collecting in March 2010. Prior to this date, we use the data obtained from Wrightson ICAP in the ELB sample. For our pre-crisis analysis, we substitute the EDR with the overnight London interbank offered rate (LIBOR), obtained from Bloomberg, because Eurodollar data are not available for that period. LIBOR is a commonly used indicator for the average rate at which banks may get short-term loans in the London interbank market, and it serves as reference rate for various debt instruments.[13] The third key rate in our analysis is a representative rate of secured funding from the repo market. We use the volume-weighted average rate for Treasury GC (general collateral) repo obtained from the FRBNY, which we will refer to as RPR. Finally, we use the overnight AA non-financial commercial paper rate (CPR) released by the Federal Reserve Board.[14] An important feature of the CPR in our context is that it represents an unsecured funding rate that is not directly affected by the changing monetary policy framework and new banking regulations discussed above.

Visual investigation suggests very strong co-movement among the rates during normal times (figure 4). Moreover, the sample means and standard deviations of the rates are remarkably close in the pre-crisis period, as shown in panel A of table 2. However, as one can infer from figure 5 and panel B of table 2, the co-movement of rates appears to have weakened over the ELB period, on net. For example, RPR remained especially elevated relative to unsecured rates around late 2011, reportedly due to longer dealer positioning in Treasury securities that coincided with the Fed's Maturity

[13]See Hou and Skeie (2014) for a detailed description of the rate-setting mechanism and efforts to reform the LIBOR.

[14]Data are available at http://www.federalreserve.gov/releases/cp/.

Figure 4. Overnight Money Market Interest Rates: 2001–07



Notes: Data are daily. EFFR and RPR are from FRBNY. LIBOR is from Bloomberg. CPR is from the commercial paper data release of the Federal Reserve Board.

Extension Program as well as higher Treasury debt issuance.[15] In addition to weaker co-movement, calendar effects relative to the level of the rates seem stronger, on average, over the ELB period, and the sample moments also show more variation across the rates. In the next section, we lay out the empirical framework to quantify such differences and analyze them in detail.

Another difference between the two samples is the degree of stationarity of the rates. In table 3, we report the augmented Dickey and Fuller (1979) (ADF) unit-root test statistic and the Elliott, Rothenberg, and Stock (1996) (ERS) point-optimal test statistic, which has higher power against persistent alternatives. In the pre-crisis sample, we cannot reject the null of a unit root in the interest rates at any conventional significance level with respect to both test statistics. In

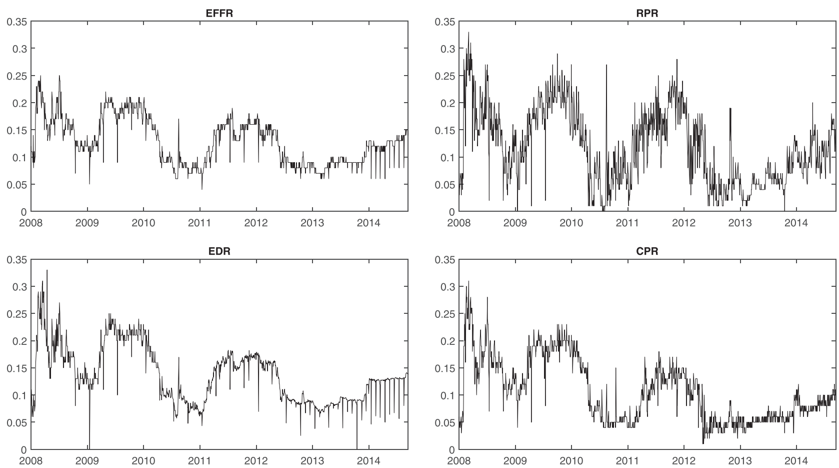
[15]During this program, the Fed sold about $650 billion of short-term securities and used the proceeds to buy longer-term securities. By extending the average maturity of its securities portfolio, the Fed aimed to put downward pressure on longer-term interest rates to ease conditions in financial markets.

Table 2. Descriptive Statistics of Money Market Rates

	EFFR	RPR	LIBOR/ EDR*	CPR
A. January 2, 2001–July 31, 2007				
Mean	2.937	2.881	2.999	2.927
St. Dev.	1.660	1.639	1.661	1.662
10th	1.010	0.980	1.058	0.990
50th	2.480	2.440	2.541	2.450
90th	5.250	5.220	5.301	5.250
AC(1)	0.999	0.999	0.999	0.999
B. December 17, 2008–August 28, 2015				
Mean	0.129	0.118	0.137	0.107
St. Dev.	0.042	0.068	0.051	0.058
10th	0.080	0.030	0.080	0.040
50th	0.130	0.110	0.130	0.090
90th	0.190	0.210	0.210	0.190
AC(1)	0.954	0.920	0.950	0.958

Notes: Data are daily. Mean standard deviation, and quantiles are reported in percent. AC(1) denotes first-order autocorrelation.
*LIBOR is used for panel A calculations and EDR is used in panel B.

Figure 5. Overnight Money Market Interest Rates: 2008–15



Notes: Data are daily. EFFR and RPR are from FRBNY. EDR is from FRBNY after March 2010, and from Wrightson ICAP prior to this date. CPR is from the commercial paper data release of the Federal Reserve Board.

Table 3. Unit-Root Tests

	EFFR	**RPR**	**LIBOR/ EDR***	**CPR**
A. ADF Test				
Pre-crisis	−1.24	−1.31	−1.00	−1.02
ELB	−3.37	−3.17	−2.82	−2.52
B. ERS Test				
Pre-crisis	251.3	275.6	158.2	195.7
ELB	1.2	2.8	2.5	3.4

Notes: ADF is the augmented Dickey and Fuller (1979) test with the 1, 5, and 10 percent critical values of −3.44, −2.87, and −2.57, respectively. ERS is the point-optimal test of Elliott, Rothenberg, and Stock (1996) with the 1, 5, and 10 percent critical values of 1.99, 3.26, and 4.48, respectively.
*LIBOR is used for panel A calculations and EDR is used in panel B.

contrast, we reject the null of unit root for all rates in the ELB sample according to the ADF test statistics, with the exception of CPR, and for all rates according to the ERS test statistic. Therefore, the interest rates are well approximated by integrated processes with a likely common stochastic trend in the pre-crisis sample, reflecting the fact that this period contains a full monetary policy cycle with easing early in the period followed by a gradual tightening beginning in 2004. In the ELB period, the rates are persistent but not integrated against the backdrop of no change in the target federal funds rate. Our modeling strategy incorporates this difference in rate dynamics.

4. Models

We specify models that account for co-movement and persistence of the money market rates as well as time variation in their volatilities and cross-correlations. We also allow for various calendar factors that likely affect dynamics of rates on specific days. The unit-root tests suggest that the interest rates are well approximated by integrated processes in the pre-crisis sample while they are persistent but stationary during the ELB sample. Therefore, we estimate different models for the pre-crisis and the ELB period.

The pre-crisis model is a vector error correction (VEC) process that incorporates the long-run equilibrium relationship of overnight money market rates. Let y_t denote the vector of the interest rates at time t, that is, $y_t = (EFFR_t, RPR_t, LIBOR_t, CPR_t)'$ in the pre-crisis sample. The interest rate dynamics are characterized by the following VEC model:

$$\Delta y_t = Ad_t + \beta \Delta TFFR_t + \sum_{j=1}^{p} \Phi_j \Delta y_{t-j} + \Theta z_{t-1} + \epsilon_t, \qquad (1)$$

where d_t is a vector of indicator variables for calendar effects, which we will explain in detail; *TFFR* is the target federal funds rate; z_t is a vector of error correction terms; and ϵ_t is a zero-mean martingale difference vector process, which is possibly heteroskedastic.[16] Reflecting the pre-crisis monetary policy operating framework, we impose the restriction that there are three distinct co-integrating relationships and that all of them involve EFFR. Formally, we have $z_{it} = y_{1t} - (c_i + \gamma_i y_{i+1,t})$, where $i = 1, 2, 3$.[17]

The vector of calendar effects, d_t, contains ten indicator variables to account for reserve maintenance period days: two indicators for elevated payment days within a month (15th and 25th), two for financial reporting days (month-end and quarter-end), and a dummy variable to control for the brief disturbance in money markets caused by the September 2001 terror attacks. As a result, the model does not contain a constant vector because it cannot be separately identified given the set of maintenance period indicators. We set $p = 4$ based on the Schwarz information criterion.[18]

There exists a mapping from this VEC system to a VAR that can be defined for the level of interest rates. This mapping allows

[16] Alternatively, one can also estimate a VAR for the level of the rates assuming stationarity, with *TFFR* on the right-hand side. We find that this alternative specification yields qualitatively the same results. The full set of results is available from the authors upon request.

[17] We obtain nearly identical results when we estimate the number of co-integrating relationships as well as the co-integration parameters in a less restricted fashion as in Johansen (1995).

[18] The total number of parameters to be estimated is equal to 140, which results in approximately forty-six observations per parameter.

us to directly compare the results from the pre-crisis period with those from the ELB period, as the model for the latter sample is a VAR in levels. Let Ψ_j for $j = 1, \ldots, p+1$ denote the autoregressive coefficient matrices in the implied VAR in levels. Then we have $\Psi_1 = \Phi_1 + I + \Theta\Gamma$, where I is an identity matrix, $\Gamma = (i, -diag\{-\gamma\})$, i is a vector of ones, γ is the vector of co-integration slopes given previously, and $diag(.)$ indicates a diagonal matrix, $\Psi_j = \Phi_j - \Phi_{j-1}$ for $j = 2, \ldots, p$, and $\Psi_{p+1} = -\Phi_p$.[19]

For the ELB period, we specify the following VAR model in levels given the stationary behavior of interest rates in this sample:

$$y_t = \Pi d_t + \sum_{j=1}^{p} \Xi_j y_{t-j} + \epsilon_t, \tag{2}$$

where d_t is now a 9×1 vector that contains month-end, quarter-end, day-of-week, and elevated payment flow-day indicators.[20] Note that the EDR replaces the LIBOR in this sample, so that $y_t = (EFFR_t, RPR_t, EDR_t, CPR_t)'$. We set $p = 3$ based on Schwarz model-selection criteria.[21]

Both visual investigation and formal testing of the model residuals suggest significant volatility clustering in both sample periods. Hence, we model volatility dynamics using multivariate GARCH models. Our modeling strategy closely follows that of Bollerslev (1990); however, instead of assuming a constant conditional correlation matrix, we allow for different correlation structures on financial reporting days. Therefore, our specification can be thought of as a hybrid of the constant correlation model and the dynamic correlation model of Engle (2002), who postulates a fully time-varying conditional correlation matrix. Let $E(\epsilon_t \epsilon_t' | \Omega_{t-1}) = H_t$, where Ω_t is the information set at time t; then we can write

$$H_t = D_t R_t D_t, \tag{3}$$

[19] A caveat is that in the pre-crisis model, shocks are permanent due to the modeling of interest rates as integrated processes.

[20] Day-of-week indicators replace those for maintenance period days, as the latter become insignificant amid abundant reserves in the ELB period.

[21] This model has eighty-four parameters to be estimated, resulting in seventy-eight observations per parameter.

where $D_t = diag\left\{\sqrt{h_{it}}\right\}$, $h_{it} = Var(\epsilon_{it}|\Omega_{t-1})$, and $R_t = Corr(\epsilon_t|\Omega_{t-1})$. The individual variances are modeled via the following GARCH specification:

$$h_{it} = \omega_i + \tau_i \epsilon_{i,t-1}^2 + \delta_i h_{i,t-1} + \lambda_{i,1} I_{m,t} + \lambda_{i,2} I_{q,t}, : i = 1, \ldots, 4, \quad (4)$$

where I_m and I_q are month-end and quarter-end indicators, respectively. In this specification, the variance at time t is essentially a weighted average of its lagged value, the new information at time $t-1$ that is captured by the most recent squared residual, the long-run unconditional variance, and the level shifts in volatility on financial reporting dates. We estimate the GARCH equation under variance targeting so that ω_i is a function of the sample variance of $\epsilon_{i,t}$ and the mean vector of the indicator series. Finally, the correlation matrix R_t is specified as follows:

$$R_t = I_{m,t} R_m + I_{q,t} R_q + (1 - I_{m,t} - I_{q,t}) R_n, \quad (5)$$

where R_m, R_q, and R_n are correlation matrices of GARCH residuals, that is, $h_{it}^{-1/2} \epsilon_t$, at month-ends, quarter-ends, and all other days, respectively.

5. Empirical Results

5.1 Co-movement of Rates and Monetary Policy Transmission

Our estimates for the pre-crisis sample are consistent with the conventional monetary policy implementation framework. As shown in panel A of table 4, lagged EFFR variables are significant in all other rate equations, implying that interest rates were adjusting in response to changes in the policy rate. Moreover, the EFFR was not responding to the other rates, as indicated by the insignificance of lagged rates in the first column. As can be seen in panel B, changes in the target federal funds rate are highly significant in all equations of the VEC model.[22] Other than the EFFR, no other interest

[22]To save space, we do not report the parameter estimates of c_i and γ_i in the co-integration equation. These estimates are highly statistically significant and very close to 0 and -1 in all equations, respectively.

Table 4. Overnight Money Market Rates before the Financial Crisis

	EFFR	**RPR**	**LIBOR**	**CPR**
	A. Autoregressive Terms (Sum)			
EFFR	0.947	0.449	0.345	0.521
	(0.00)	(0.00)	(0.00)	(0.00)
RPR	0.033	0.694	0.010	0.001
	(0.39)	(0.00)	(0.72)	(0.98)
LIBOR	0.016	−0.125	0.546	−0.091
	(0.91)	(0.41)	(0.00)	(0.43)
CPR	0.005	−0.021	0.099	0.570
	(0.97)	(0.92)	(0.21)	(0.00)
	B. Other Variables			
ΔTFFR	0.454	0.406	0.337	0.416
	(0.00)	(0.00)	(0.00)	(0.00)
15th	5.50	6.04	6.10	7.00
	(0.00)	(0.00)	(0.00)	(0.00)
25th	4.33	0.69	0.09	1.14
	(0.00)	(0.24)	(0.75)	(0.00)
Month-End	5.33	4.15	7.43	6.20
	(0.00)	(0.00)	(0.00)	(0.00)
Quarter-End	5.66	−12.52	17.33	11.17
	(0.03)	(0.01)	(0.00)	(0.00)

Notes: Columns represent equations of the models. The sum of autoregressive terms correspond to $\Sigma\Psi_j$ in terms of the notation of section 4. p-values based on robust (HAC) standard errors are reported in parentheses. Calendar effects are in *basis points*. Daily sample runs from January 2, 2001, to July 31, 2007.

rate in the system had predictive power for the remaining interest rates. Overall, these results confirm that overnight funding rates were adjusting in response to policy intervention and dynamics in the federal funds market. These results are consistent with the view that the overnight money markets were tightly connected through the federal funds market in the pre-crisis period.

The estimates from the ELB sample shown in panel A of table 5 paint a different picture. The federal funds and Eurodollar markets appear to be closely connected, as indicated by the statistical and

Table 5. Overnight Money Market Rates at the ELB

	EFFR	RPR	EDR	CPR
A. Autoregressive Terms (Sum)				
EFFR	0.911	0.107	0.223	0.153
	(0.00)	(0.21)	(0.00)	(0.02)
RPR	0.032	0.809	0.14	−0.011
	(0.00)	(0.00)	(0.29)	(0.47)
EDR	−0.024	0.048	0.705	0.000
	(0.53)	(0.50)	(0.00)	(1.00)
CPR	0.036	0.054	0.069	0.881
	(0.01)	(0.14)	(0.00)	(0.00)
B. Other Variables				
15th	0.80	3.29	0.85	0.96
	(0.00)	(0.00)	(0.00)	(0.00)
25th	−0.26	0.65	−0.08	0.37
	(0.01)	(0.08)	(0.36)	(0.05)
Month-End	−0.14	3.47	−0.13	0.37
	(0.63)	(0.00)	(0.72)	(0.20)
Quarter-End	−3.21	−0.41	−5.07	−1.58
	(0.00)	(0.70)	(0.00)	(0.03)

Notes: Columns represent equations of the models. The sum of autoregressive terms correspond to $\Sigma\Xi_j$ in terms of the notation of section 4. p-values based on robust (HAC) standard errors are reported in parentheses. Calendar effects are in *basis points*. Daily sample runs from December 17, 2008, to August 28, 2015.

economic significance of the EFFR coefficients in the EDR equation. Similarly, the EFFR is linked to the CPR, which is another key unsecured rate in the system, although to a lesser extent than the EDR. These results imply that the EFFR continued to be an anchor for unsecured rates in the ELB period, although its transmission weakened relative to pre-crisis norms, especially in the case of the CPR.

The most dramatic change across the two periods concerns the transmission from the federal funds to the repo market. The EFFR no longer predicts RPR movements in the ELB period. This disconnect between the two rates that used to move in tandem also

Table 6. Overnight Money Market Rates before the ON RRP

	EFFR	**RPR**	**EDR**	**CPR**
A. Autoregressive Terms (Sum)				
EFFR	0.911	0.112	0.226	0.164
	(0.00)	(0.26)	(0.00)	(0.03)
RPR	0.018	0.803	0.002	−0.022
	(0.13)	(0.00)	(0.88)	(0.23)
EDR	−0.002	0.045	0.739	0.005
	(0.97)	(0.59)	(0.00)	(0.94)
CPR	0.028	0.052	0.047	0.880
	(0.04)	(0.20)	(0.01)	(0.00)
B. Calendar Effects				
15th	1.11	3.93	1.26	1.42
	(0.00)	(0.00)	(0.00)	(0.00)
25th	−0.29	0.70	0.01	0.54
	(0.03)	(0.17)	(0.94)	(0.02)
Month-End	0.79	3.94	1.19	0.71
	(0.00)	(0.00)	(0.00)	(0.06)
Quarter-End	−3.14	−1.06	−4.29	−2.02
	(0.00)	(0.45)	(0.00)	(0.04)

Notes: Columns represent equations of the models. The sum of autoregressive terms correspond to $\Sigma\Xi_j$ in the notation of section 4. p-values based on robust (HAC) standard errors are reported in parentheses. Calendar effects are in *basis points*. Daily sample runs from December 17, 2008, to September 20, 2013.

emphasizes the diminished role of banks as arbitrageurs, as discussed in section 2.1. Overall, we conclude that co-movement of the EFFR with other rates became noticeably weaker in the ELB sample amid superabundant reserves, subdued trading, and dominance of IOR arbitrage trades in the federal funds market.

To assess the effects of the ON RRP facility on money market dynamics, we now focus solely on the ELB period and estimate VAR models for the two subsamples separated by the inception of the ON RRP facility in September 2013. The comparison of the results summarized in tables 6 and 7 suggests that the ON RRP had two

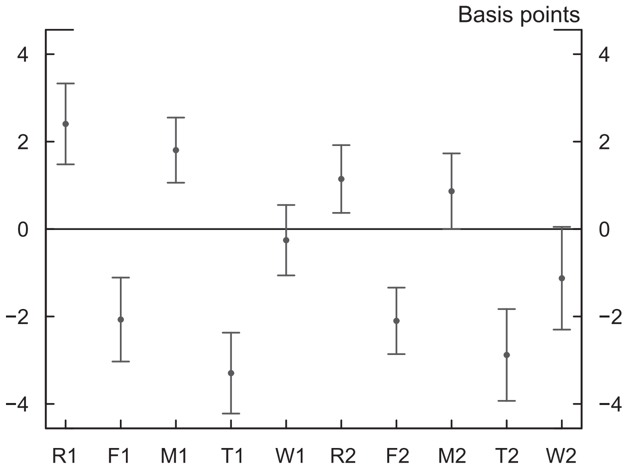
Table 7. Overnight Money Market Rates after the ON RRP

	EFFR	**RPR**	**EDR**	**CPR**
A. Autoregressive Terms (Sum)				
EFFR	0.823	0.033	0.290	0.333
	(0.00)	(0.83)	(0.06)	(0.00)
RPR	0.102	0.813	0.096	0.084
	(0.00)	(0.00)	(0.00)	(0.00)
EDR	−0.127	0.101	0.416	−0.145
	(0.01)	(0.42)	(0.00)	(0.01)
CPR	0.129	0.126	0.113	0.565
	(0.04)	(0.21)	(0.12)	(0.00)
B. Calendar Effects				
15th	−0.02	1.61	−0.18	−0.24
	(0.87)	(0.00)	(0.28)	(0.27)
25th	−0.08	0.46	−0.22	−0.12
	(0.50)	(0.17)	(0.09)	(0.54)
Month-End	−2.35	2.42	−3.25	−0.37
	(0.00)	(0.00)	(0.00)	(0.15)
Quarter-End	−3.41	1.10	−6.80	−1.05
	(0.00)	(0.37)	(0.00)	(0.02)

Notes: Columns represent equations of the models. The sum of autoregressive terms correspond to $\Sigma\Xi_j$ in the notation of section 4. p-values based on robust (HAC) standard errors are reported in parentheses. Calendar effects are in *basis points*. Daily sample runs from September 23, 2013, to August 28, 2015.

important effects. First, transmission from the EFFR to the other unsecured rates clearly improved: the sum of lagged EFFR terms increased from 0.23 to 0.29 in the case of the EDR and from 0.16 and 0.33 in the case of the CPR. Second, the RPR became a significant predictor of the EFFR movements, in contrast to the precrisis relationship where RPR was moving in response to changes in the EFFR. Interestingly, RPR became highly significant in the EDR and CPR equations, emphasizing the growing importance of the repo market. Hence, it appears that the ON RRP markedly improved the overall co-movement of overnight interest rates and

Figure 6. Day of Maintenance Period Effects on EFFR during the Pre-crisis Period



Notes: Dots indicate point estimates and horizontal lines mark the boundaries of the 95 percent confidence bands. M, T, W, R, and F denote days of the week from Monday to Friday. The subscripts indicate whether the corresponding date is the first or the second one in the maintenance period.

transmission from the federal funds market to other segments of unsecured funding markets.

5.2 Reserve Maintenance Period Effects

In figure 6, we report point and interval estimates for the coefficients of the effects of reserve maintenance days on the EFFR in the pre-crisis period. Different days of the week had small but economically meaningful and statistically significant effects on the EFFR. For example, due to elevated payment flows following weekends, the EFFR used to be firmer by 1 to 2 basis points on Mondays. By contrast, funds used to trade softer by a slightly greater magnitude on Fridays, as banks generally tried to avoid an excess position over the weekend, during which reserves count for three days toward the reserve requirement. Tuesdays were also associated with softness due to reduced demand towards the middle of the week when payment flows are relatively lighter. These estimates are consistent with those of Hamilton (1996), Carpenter and Demiralp (2006), and Judson and Klee (2010) that were obtained in different empirical frameworks.

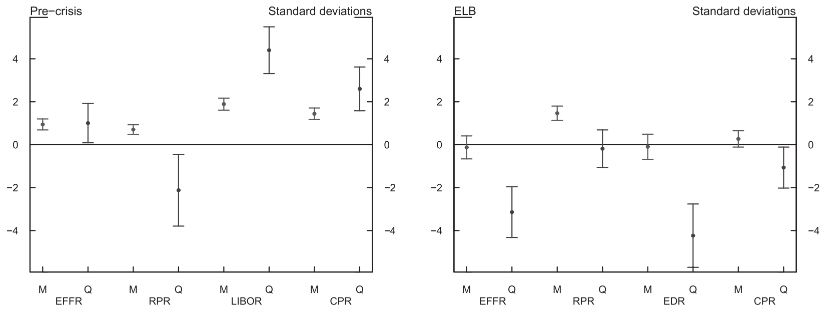
In the ELB period, although we cannot statistically reject day-of-the-week effects in the federal funds market, our estimates (not reported) imply miniscule effects. When we combine our coefficient estimates with trading volumes reported by Afonso, Entz, and LeSueur (2013b), we find that the average day-of-week effect is about only about 3 percent of its pre-crisis level in dollar terms. Moreover, when we normalize the estimated effects by the standard deviation of the EFFR residuals to control for the dramatically different level of the average EFFR across the two periods, we find that the day-of-week effect is about 70 percent weaker in the ELB period. Therefore, given the abundance of reserves and their fairly widespread distribution as reported by Ennis and Wolman (2015), we conclude that reserve maintenance effects in the federal funds market diminished substantially.

5.3 Market Dynamics on Financial Reporting Days

The estimated magnitudes of calendar effects are quite different across the pre-crisis and ELB periods, as is evident from panel B in tables 4 and 5. However, the average levels of overnight interest rates are also dramatically different across the two samples. To control for the general level of interest rates and allow for a direct comparison between the two periods, we normalize the estimates relative to standard deviations of model residuals from the respective equation in the VAR system.

Figure 7 shows the *normalized* estimates for the two main samples. In the pre-crisis sample, all rates were subject to modest upward pressure at month-ends, possibly due to heavier payment flows as well as adjustments to balance sheets related to financial reporting. Most comprehensive financial reports are produced on a quarterly basis, so deleveraging by financial intermediaries on quarter-end is quite common. Indeed, quarter-end effects were more prominent than month-end effects, with the exception of the EFFR. Rates were markedly softer in the repo market, likely because of lower financing demand from dealers managing their leverage. In contrast, it appears that reduced willingness to lend in unsecured markets on quarter-ends was the dominant factor leading to higher rates on financial

Figure 7. Month- and Quarter-End Effects

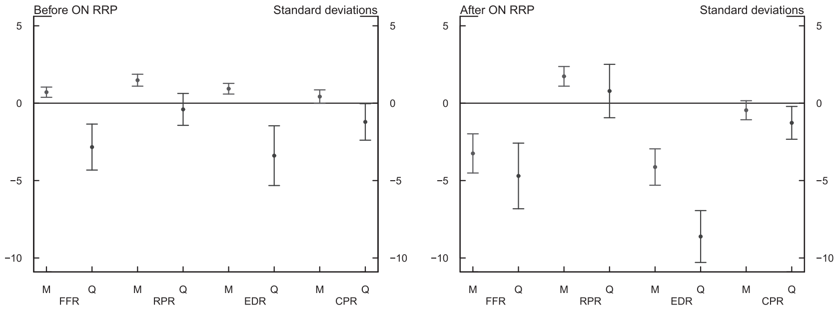


Notes: Dots indicate point estimates and horizontal lines mark the boundaries of the 95 percent confidence bands. M and Q denote month-end and quarter-end, respectively. Effects are normalized with respect to the standard deviations of model residuals.

reporting days. This pattern is observed especially for LIBOR, likely reflecting banks' desire to show strong liquidity positions on their financial statements and regulatory filings.

Money market dynamics on financial reporting days changed materially in the ELB sample. First of all, both the EFFR and the EDR started softening on quarter-ends, reflecting fewer IOR-arbitrage trades by foreign banks and large domestic banks. Balance sheet constraints associated with the new FDIC assessment scheme and Basel III leverage ratio that became prevalent in the later part of the ELB sample largely explain the reduced borrowing demand on quarter-ends. Contrary to the case of unsecured rates, the quarter-end effect became insignificant in the repo market, on net, likely reflecting a combination of factors. First, earlier in the ELB period, collateral demand was relatively strong due to flight-to-quality flows, leading to increased willingness to lend cash at lower rates in lieu of Treasury collateral. Second, later in the period, as new regulations were announced and implemented, lower dealer leverage and reduced net repo financing reduced the scope of quarter-end deleveraging effects. Finally, the availability of the ON RRP as a viable investment, especially on financial reporting dates when other investment options may be limited, reduced the potential for sharp falls in the repo rates. In terms of the CPR, cash lenders' search for alternative investments on quarter-ends amid weaker

Figure 8. Month- and Quarter-End Effects within the ELB Period



Notes: Dots indicate point estimates and horizontal lines mark the boundaries of the 95 percent confidence bands. M and Q denote month-end and quarter-end, respectively. Effects are normalized with respect to the standard deviations of model residuals.

demand by bank borrowers appears to have led to a softening in this rate.

Figure 8 shows the normalized calendar effects on rates through the ELB period. Both the EFFR and the EDR started to decline notably at month-ends later in the ELB sample as foreign banks withdrew from the market for financial reporting purposes. Moreover, downward pressure on these rates at quarter-ends also became more pronounced, especially for the EDR. This likely reflects the fact that Eurodollars are a relatively more important source of dollar funding for foreign banks, which are subject to less stringent implementation of the Basel III leverage ratio. In contrast, quarter-end effects on CPR have been relatively stable, suggesting limited spillover effects from the federal funds and Eurodollar markets. The absence of direct implications on the non-financial commercial paper market also suggests that the leverage ratio requirements have indeed been the primary driver for unsecured rates on financial reporting days.

5.4 Volatility and Correlation of Overnight Interest Rates

We now focus on both general and financial-reporting-driven volatility dynamics across the two main sample periods as well as before

Table 8. Volatility of Rates

	Pre-crisis				ELB			
	EFFR	**RPR**	**LIBOR**	**CPR**	**EFFR**	**RPR**	**EDR**	**CPR**
σ_ε	5.64	5.91	3.94	4.30	1.05	2.38	1.22	1.51
τ	0.116	0.306	0.450	0.316	0.253	0.231	0.440	0.414
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
δ	0.839	0.173	0.180	0.171	0.396	0.385	0.240	0.282
	(0.00)	(0.11)	(0.03)	(0.01)	(0.05)	(0.00)	(0.00)	(0.00)

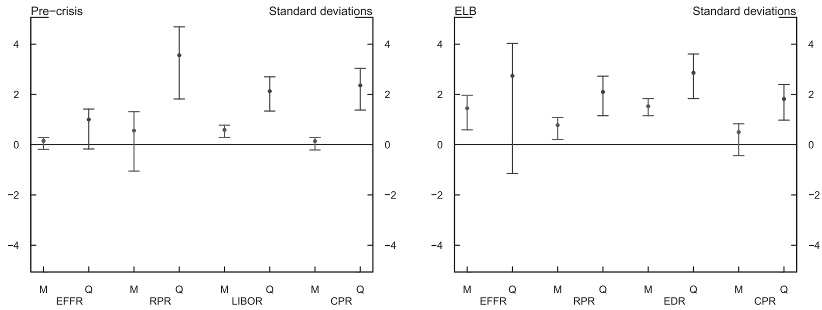
Notes: p-values based on robust standard errors are reported in parentheses. σ_ε are reported in *basis points.*

and after the inception of the ON RRP facility. The parameter estimates of the volatility models for the pre-crisis and ELB samples are shown in table 8. As expected, volatility of all rates declined substantially at the ELB in absolute terms. For example, the volatility of innovations in the EFFR equation declined from 5.6 basis points to only about 1 basis point. Meanwhile, the volatility process for the EFFR has become notably less persistent, as captured by the decline in the sum of GARCH parameters ($\tau + \delta$), and more responsive to shocks, as measured by the increase in the coefficient of the squared innovation (τ). Therefore, aside from calendar effects, which will be discussed in more detail below, volatility clustering in the EFFR became prevalent amid subdued trading activity in the federal funds market. Meanwhile, the volatility of the repo rate has become somewhat more persistent, with a slight increase in sensitivity to shocks.

Figure 9 shows the estimated month-end and quarter-end effects on volatilities in the pre-crisis and the ELB samples.[23] As before, estimates are normalized by dividing by the standard deviations of residuals to allow for direct comparison across the two periods. Prior to the crisis, similar to the calendar effects in the conditional mean models, quarter-ends had larger effects on rate volatilities

[23]The confidence bands that are based on asymptotic normal distribution are asymmetric, as they are estimated in the variance space and then converted to standard deviations.

Figure 9. Month- and Quarter-End Effects on Volatility



Notes: Dots indicate point estimates and horizontal lines mark the boundaries of the 95 percent confidence bands. M and Q denote month-end and quarter-end, respectively. Effects are normalized with respect to the standard deviations of model residuals.

Table 9. Volatility of Rates within the ELB Period

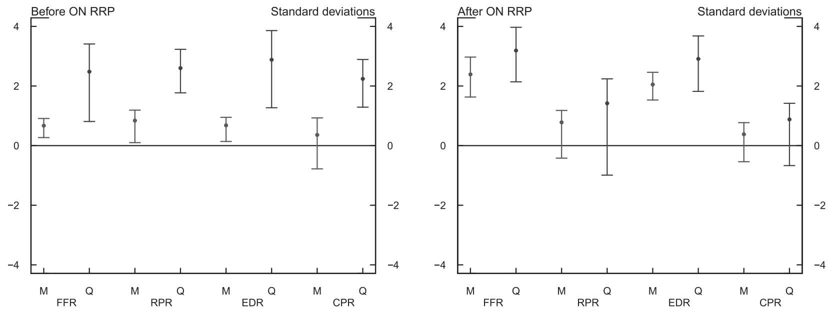
	Before ON RRP				After ON RRP			
	EFFR	**RPR**	**EDR**	**CPR**	**EFFR**	**RPR**	**EDR**	**CPR**
σ_ε	1.11	2.65	1.27	1.67	0.73	1.40	0.79	0.83
τ	0.212	0.159	0.365	0.383	0.189	0.315	0.458	0.158
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
δ	0.561	0.327	0.368	0.281	0.191	0.465	0.146	0.681
	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.01)	(0.00)	(0.00)

Notes: p-values based on robust standard errors are reported in parentheses. σ_ε are reported in *basis points.*

than month-ends. Most notably, RPR exhibited substantial volatility clustering with two to five times higher volatility on quarter-ends than other times. In the ELB period, quarter-end volatility in the RPR moderated significantly, while becoming higher in the case of the EFFR.

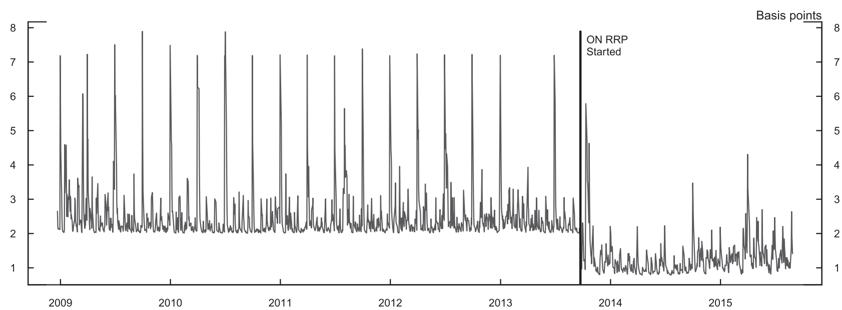
Consistent with the soft floor set by the ON RRP facility, volatility of the overnight interest rates declined 35 to 50 percent in the second half of the ELB sample, as shown in table 9. Moreover, the estimated parameters indicate a substantial reduction in volatility clustering of the RPR, mainly led by a dramatic decline in the calendar effects (table 9 and figure 10). Figure 11 illustrates this striking

Figure 10. Month- and Quarter-End Effects on Volatility with the ELB Period



Notes: Dots indicate point estimates and horizontal lines mark the boundaries of the 95 percent confidence bands. M and Q denote month-end and quarter-end, respectively. Effects are normalized with respect to the standard deviations of model residuals.

Figure 11. Repo Rate Volatility and ON RRP



change in the volatility of RPR following the inception of the ON RRP.[24] Elevated-volatility episodes in the fall of 2013 are related to the Treasury's debt limit impasse and government shutdown. We also observe that calendar effects on RPR volatility largely diminished, likely reflecting the fact that the facility rate limiting the

[24]One caveat in these estimates is that the unconditional variances of the GARCH specifications are anchored to the corresponding sample variances. Therefore, the level shift after the ON RRP inception reflects the *average* effect across the two ELB subsamples.

Table 10. Correlations of VAR Residuals

	Pre-crisis			ELB		
	RPR	**LIBOR**	**CPR**	**RPR**	**EDR**	**CPR**
Normal Times	0.490	0.586	0.614	0.457	0.545	0.373
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Month-End	0.421	0.246	0.341	0.301	0.879	0.395
	(0.04)	(0.47)	(0.22)	(0.19)	(0.00)	(0.37)
Quarter-End	0.348	0.334	0.362	−0.056	0.564	0.360
	(0.30)	(0.23)	(0.32)	(1.00)	(0.03)	(0.29)

Notes: Correlations with EFFR. p-values based on robust standard errors are reported in parentheses.

potential for sharp falls in rates on financial reporting dates when investment options are limited for cash lenders in the repo market.

Similar to the case of the RPR, the quarter-end effect on the CPR also became insignificant in the second half of the ELB sample. In contrast, month-end and quarter-end effects became more pronounced for other unsecured rates, mainly due to the pronounced pullback from the federal funds and Eurodollar markets driven by the Basel III regulations.

Correlation structure of VAR innovations provides further insights into the co-movement of overnight interest rates. Table 10 reports estimates for the pre-crisis and ELB samples from the multivariate GARCH framework. Interestingly, the correlations of the EFFR residuals with those of the other three rates during normal times are fairly close across the two main samples. Hence, it appears that factors exogenous to the dynamic system of these rates, such as Treasury debt issuance and related liquidity effects, continued to operate in a similar fashion, on net. We also compare correlations within the ELB sample and report the results in table 11. Although reliable comparisons are not possible in some cases given the limited number of month-end and quarter-end observations in the ELB subsamples, it appears that the correlation of EFFR and RRP innovations declined over time. Meanwhile, the EFFR innovations became more strongly correlated with those of the EDR. This increased correlation is more pronounced on

Table 11. Correlations of VAR Residuals within the ELB Period

	Before ON RRP			After ON RRP		
	RPR	**EDR**	**CPR**	**RPR**	**EDR**	**CPR**
Normal Times	0.502	0.546	0.413	0.128	0.612	0.173
	(0.00)	(0.00)	(0.00)	(0.16)	(0.00)	(0.07)
Month-End	0.395	0.596	0.104	−0.291	0.854	0.039
	(0.17)	(0.05)	(0.63)	(1.00)	(0.00)	(0.90)
Quarter-End	0.032	0.595	0.358	−0.489	0.356	0.334
	(0.95)	(0.05)	(0.34)	(1.00)	(0.59)	(0.51)

Notes: Correlations with EFFR. p-values based on robust standard errors are reported in parentheses.

month-ends, likely reflecting the effects of the Basel III leverage ratio requirements.

6. Conclusion

We analyze changing dynamics of overnight money markets in the wake of the global financial crisis and the resulting policy response. To that end, we model long-run and short-run dynamics of a set of key money market rates as well as their interrelations in a multivariate framework. We analyze the effects of changing monetary and regulatory policy that evolved in the aftermath of the crisis on money markets.

We find that co-movement of money market rates weakened in the ELB period compared with the historical norms. Although the federal funds rate continued to provide an anchor for unsecured overnight interest rates, its transmission to the repo market was hampered. Moreover, the day-of-week effects on the federal funds rate have substantially diminished, reflecting the abundance of bank reserves and their fairly widespread distribution.

New banking regulations and the Fed's ON RRP facility introduced in 2013 have further transformed the money markets. Movements in unsecured short-term funding costs around financial reporting days have been exacerbated by increased costs of large balance sheets in the new regulatory environment. Consistent with the intended effect of the ON RRP to set a soft floor for overnight funding rates, interest rate co-movement improved and rate volatilities, especially in the repo market, substantially declined after the ON RRP operations started. Moreover, calendar effects in the repo market largely disappeared, reflecting diminished potential for drops in rates, as well as the availability of reverse repos with the Fed as a viable investment option around financial reporting days when other alternatives are limited.

References

Adrian, T., M. Fleming, J. Goldberg, M. Lewis, F. M. Natalucci, and J. J. Wu. 2013. "Dealer Balance Sheet Capacity and Market Liquidity During the 2013 Selloff in Fixed Income Markets." FEDS Notes (October 16). Board of Governors of the Federal Reserve System.

Adrian, T., and H. S. Shin. 2010. "Liquidity and Leverage." *Journal of Financial Intermediation* 19 (3): 418–37.

Afonso, G., A. Entz, and E. LeSueur. 2013a. "Who's Borrowing in the Fed Funds Market." Liberty Street Economics Blog, December 9. Available at http://libertystreeteconomics.newyorkfed.org/2013/12/whos-borrowing-in-the-fed-funds-market.html.

———. 2013b. "Who's Lending in the Fed Funds Market." Liberty Street Economics Blog, December 2. Available at http://libertystreeteconomics.newyorkfed.org/2013/12/whos-lending-in-the-fed-funds-market.html.

Afonso, G., A. Kovner, and A. Schoar. 2011. "Stressed, Not Frozen: The Federal Funds Market in the Financial Crisis." *Journal of Finance* 66 (4): 1109–39.

Anbil, S., and Z. Senyuz. 2016. "Window-Dressing and Trading Dynamics in the Triparty Repo Market." SSRN Working Paper.

Anderson, A. G., J. E. Ihrig, G. C. Weinbach, and E. E. Meade. 2016. "What Happened in Money Markets After the Fed's December Rate Increase?" FEDS Notes (February 22). Board of Governors of the Federal Reserve System. Available at http://dx.doi.org/10.17016/2380-7172.1713.

Ashcraft, A. B., M. L. Bech, and W. S. Frame. 2010. "The Federal Home Loan Bank System: The Lender of Next-to-Last Resort?" *Journal of Money, Credit and Banking* 42 (4): 551–83.

Baklanova, V., C. Caglio, M. Cipriani, and A. Copeland. 2016. "A New Survey of the U.S. Bilateral Repo Market: A Snapshot of Broker-Dealer Activity." Staff Report No. 758, Federal Reserve Bank of New York.

Bech, M., E. Klee, and V. Stebunovs. 2014. "Arbitrage, Liquidity and Exit: The Repo and Federal Funds Markets Before, During, and Emerging from the Financial Crisis." In *Developments in Macro-Finance Yield Curve Modelling*, ed. J. S. Chadha, A. C. J. Durré, M. A. S. Joyce, and L. Sarno, 293–325 (chapter 11). Cambridge: Cambridge University Press.

Bollerslev, T. 1990. "Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model." *Review of Economics and Statistics* 72 (3): 498–505.

Carpenter, S., and S. Demiralp. 2006. "The Liquidity Effect in the Federal Funds Market: Evidence from Daily Open Market Operations." *Journal of Money, Credit and Banking* 38 (June): 901–20.

Choi, D. B., and H.-S. Choi. 2016. "The Effect of Monetary Policy on Bank Wholesale Funding." Staff Report No. 759, Federal Reserve Bank of New York.

Cipriani, M., and J. Gouny. 2015. "The Eurodollar Market in the United States." Liberty Street Economics Blog, May 27. Available at http://libertystreeteconomics.newyorkfed.org/2015/05/the-eurodollar-market-in-the-united-states.html#.V6OE7KIsC1w.

Copeland, A., I. Davis, E. LeSueur, and A. Martin. 2014. "Lifting the Veil on the U.S. Bilateral Repo Market." Liberty Street Economics Blog, July 9. Available at http://libertystreeteconomics.newyorkfed.org/2014/07/lifting-the-veil-on-the-us-bilateral-repo-market.html#.VzowDuTYgTV.

Copeland, A., A. Martin, and M. Walker. 2014. "Repo Runs: Evidence from the Tri-Party Repo Market." *Journal of Finance* 69 (6): 2343–80.

D'Amico, S., W. English, D. López-Salido, and E. Nelson. 2012. "The Federal Reserve's Large-scale Asset Purchase Programmes: Rationale and Effects." *Economic Journal* 122 (564): F415–F446.

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.

Elliott, G., T. J. Rothenberg, and J. H. Stock. 1996. "Efficient Tests for an Autoregressive Unit Root." *Econometrica* 64 (4): 813–36.

Engle, R. 2002. "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models." *Journal of Business & Economic Statistics* 20 (3): 339–50.

Ennis, H. M., and A. L. Wolman. 2015. "Large Excess Reserves in the United States: A View from the Cross-Section of Banks." *International Journal of Central Banking* 11 (1, January): 251–87.

Frost, J., L. Logan, A. Martin, P. McCabe, F. Natalucci, and J. Remache. 2015. "Overnight RRP Operations as a Monetary Policy Tool: Some Design Considerations." Finance and Economics Discussion Series No. 2015-010, Board of Governors of the Federal Reserve System.

Gorton, G., and A. Metrick. 2012. "Securitized Banking and the Run on Repo." *Journal of Financial Economics* 104 (3): 425–51.

Griffiths, M. D., and D. B. Winters. 1995. "Day-of-the-Week Effects in Federal Funds Rates: Further Empirical Findings." *Journal of Banking & Finance* 19 (7): 1265–84.

Hamilton, J. D. 1996. "The Daily Market for Federal Funds." *Journal of Political Economy* 104 (1): 26–56.

Hou, D., and D. Skeie. 2014. "LIBOR: Origins, Economics, Crisis, Scandal, and Reform." Staff Report No. 667, Federal Reserve Bank of New York.

Ihrig, J. E., G. C. Weinbach, and E. E. Meade. 2015. "Rewriting Monetary Policy 101: What's the Fed's Preferred Post-Crisis Approach to Raising Interest Rates?" *Journal of Economic Perspectives* 29 (4): 177–98.

Johansen, S. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Judson, R., and E. Klee. 2010. "Whither the Liquidity Effect: The Impact of Federal Reserve Open Market Operations in Recent Years." *Journal of Macroeconomics* 32 (3): 713–31.

Kreicher, L. L., R. N. McCauley, and P. McGuire. 2013. "The 2011 FDIC Assessment on Banks' Managed Liabilities: Interest Rate and Balance Sheet Responses." BIS Working Paper No. 413.

Krishnamurthy, A., and A. Vissing-Jorgensen. 2011. "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy." *Brookings Papers on Economic Activity* 2011 (Fall): 215–87.

Kroeger, A., and A. Sarkar. 2016. "Monetary Policy Transmission Before and After the Crisis." Liberty Street Economics Blog, June 16. Available at http://libertystreeteconomics.newyorkfed.org/2016/06/monetary-policy-transmission-before-and-after-the-crisis.html#.V6OF3KIsC1w.

Munyan, B. 2015. "Regulatory Arbitrage in Repo Markets." Working Paper No. 15-22, Office of Financial Research.

Potter, S. 2015. "Money Markets and Monetary Policy Normalization." Remarks at the Money Marketeers of New York University, New York City, April 15. Available at https://www.newyorkfed.org/newsevents/speeches/2015/pot150415.html.

Spindt, P. A., and J. R. Hoffmeister. 1988. "The Micromechanics of the Federal Funds Market: Implications for Day-of-the-Week Effects in Funds Rate Variability." *Journal of Financial and Quantitative Analysis* 23 (4): 401–16.

Yoldas, E., and Z. Senyuz. 2015. "Financial Stress and Equilibrium Dynamics in Money Markets." Finance and Economics Discussion Series No. 2015-091, Board of Governors of the Federal Reserve System.

The Determinants of Credit Union Failure: Insights from the United Kingdom*

Jamie Coen,[a,b] William B. Francis,[a] and May Rostom[c]
[a]Bank of England
[b]London School of Economics
[c]Bank of Canada

We use a proprietary data set on credit unions from the United Kingdom to develop an early-warning model of credit union failure. We find that a small set of financial attributes related to capital adequacy, asset quality, earnings performance, and liquidity, augmented with unemployment rates, reliably identifies failures within one year. Our results support the existing literature which, to date, has largely relied on data from the United States. This work therefore provides further evidence for establishing early-warning criteria for use by international regulators.

JEL Codes: G21, G28, G38.

1. Introduction

Credit unions provide an important source of credit and an avenue for saving in many countries around the world. At the end of 2016, over 68,000 credit unions located in 109 countries were serving 235 million people, or roughly 15 percent of economically active adults worldwide.[1] Since the mid-1990s, the global reach of these institutions has more than doubled and their assets have increased almost fourfold to $1.8 trillion (roughly 2 percent of world GDP). The

*The views expressed in this paper are those of the authors, and not necessarily those of the Bank of England or its committees. We are grateful to Marty Bonus, Stephen Dickinson, Chris Donald, Marcela Hashim, Sujit Kapadia, Luc Laeven, Katharine Neiss, Rhiannon Sowerbutts, and an anonymous referee as well as participants at the Bank of England's One Bank Research Seminar and Research Away Day for helpful comments. All remaining errors are ours.

[1]See World Council of Credit Unions Statistical Reports (available at http://www.woccu.org).

presence of credit unions is even more pronounced in several countries, including Ireland, Canada, and the United States, where these institutions reach more than a half of economically active adults.

In many countries, credit unions operate in rural or regional markets where access to traditional banks is more limited, thereby supporting economic activity that would otherwise go unfunded. Research has also shown that the presence of these institutions in general has implications for the real economy. For example, their presence leads commercial banks to offer higher rates on retail deposits (e.g., Hannan 2002) and lower rates on consumer loans (e.g., Feinberg 2003). These studies support the view that competition from credit unions can have direct benefits for consumers and the economy more broadly.[2] Therefore, their failure can have adverse effects not only on small, local pockets within a country but also on the wider economy to the extent that such failures affect competition. Understanding the early-warning signs of failure can therefore be important for regulators to avoid detrimental economic effects.

The literature examining the determinants of failure in the commercial banking sector is relatively widespread.[3] However, there is much less work on this issue for the credit union sector, especially outside the United States. One reason for this is because credit union failures in other countries are relatively rare.[4] Another explanation is that credit unions are less exposed to business cycle swings and

[2]Feinberg and Meade (2017) estimate the value of these direct benefits at $16 billion per year for the United States. Using a full general equilibrium analysis, they quantify the annual loss to GDP ($14.2 billion) and jobs (88,000) that result from a material (50 percent) reduction in the number of credit unions and competitive pressure on market rates that ensues. Their results have implications for understanding the economic impacts from a decline in credit unions and the merits of policies directed at fostering competition in banking markets more widely.

[3]Indeed, the large number of failures that occurred in the United States during the banking crisis in the late 1980s and early 1990s spawned an extensive body of work on the determinants of bank failure (e.g., Demirguc-Kunt 1989; Whalen 1991; Cole, Cornyn, and Gunther 1995; Cole and Gunther 1995; Sahajwala and Van den Bergh 2000; Wheelock and Wilson 2000). Studies investigating the determinants of bank failure during the 2007–09 financial crisis confirm that many of the factors explaining failure in the earlier crisis also contributed to failure in the more recent crisis (e.g., see Cole and White 2012).

[4]Rather than let credit unions fail, most jurisdictions have tended to transfer troubled credit unions to healthier ones (e.g., Jones 2010; McKillop, Ward, and Wilson 2011).

are better able to withstand shocks to their balance sheets (e.g., Smith and Woodbury 2010).[5]

Using regulatory data on credit unions in the United Kingdom, we investigate the determinants of failure within the U.K. credit union sector, which, to our knowledge, has not been examined previously. We adopt the framework employed in many of the aforementioned studies on commercial bank failure and use CAMEL factors (capital adequacy, asset quality, management, earnings performance, and liquidity) to develop an early-warning model of credit unions that failed between 2003 and 2015. Recent studies evaluating distress in mutually owned institutions, analogous to U.K. credit unions, employ similar techniques and find that CAMEL factors are helpful in identifying trouble early (Fiordelisi and Mare 2013; Francis 2014; Mare 2015). While our analysis focuses on the United Kingdom, it has implications for supervisory oversight in other countries where credit unions have a material presence, such as Australia, Canada, and Ireland.[6]

Despite playing a smaller role than traditional banking institutions in providing credit and deposit services, U.K. credit unions remain firmly on the radar of prudential supervisors given the important role they play in supporting local economies.[7] They do this mainly by supplying credit to consumers who typically find it difficult to borrow from traditional financial institutions, thus fostering economic activity that might otherwise go unfunded (Jones 2016).[8]

---

[5] On the other hand, in the United States, where credit union failures have been more common, studies find that both macroeconomic and firm-specific factors have contributed to the demise of several thousand federally insured credit unions (e.g., see Wilcox 2005, 2007). These studies also show that the failure rates and the underlying drivers for these institutions' demise were not dissimilar to those for small, federally insured commercial banks in the United States.

[6] At the end of 2016, credit unions in Australia, Canada, and Ireland were serving approximately 20 percent, 50 percent, and 75 percent of economically active adults, respectively. See the World Council of Credit Unions 2016 Statistical Report (available at http://www.woccu.org/documents/2016_Statistical_Report).

[7] The Bank of England's Prudential Regulation Authority (PRA) supervises roughly 500 credit unions with close to two million members and £3 billion in assets. Monitoring the health of individual credit unions and the sector overall is part of the PRA's overall remit and is consistent with its primary safety and soundness objective.

[8] As discussed below, U.K. credit unions are also restricted as to the amount and duration of loans they can make to one borrower. This means that the vast

In the United Kingdom the credit union sector continues to garner attention in light of government efforts to widen financial inclusion and promote effective competition in the banking sector (e.g., see Hope 2010; Jones 2016), further motivating our U.K. focus.

Our paper evaluates the characteristics of failed credit unions one year prior to their demise.[9] To define credit union failure, we use a data set on institutions that have been referred to the U.K. Financial Services Compensation Scheme (FSCS) for depositor payout as part of the formal administration process.[10] We combine this information with annual regulatory report data to create a comprehensive database of credit union failures and financial statement information covering Great Britain and Northern Ireland for the period 2002 to 2015. We supplement this data set with unemployment rates to examine their contribution to credit union condition.

Consistent with the bank failure literature, we find that a small set of financial attributes related to capital adequacy, asset quality, earnings performance, and liquidity is significant in explaining credit union failure. Our analysis highlights the relative importance of each feature in explaining credit union failure. While capital adequacy plays a prominent role, our contribution is that we also show that the other factors, including the proportion of unsecured loans as well as national and regional unemployment rates, are important for characterizing failure. Out-of-sample performance results indicate that this parsimonious set of firm-level characteristics, along with unemployment measures, classifies failures reliably while keeping false positive (type II) error rates at modest levels. Overall, these

majority make only small consumer loans and not residential mortgages, which have historically been a particular source of risk in the commercial banking sector (e.g., Antoniadis 2017).

[9] Our focus on a one-year forecast horizon is consistent with standard practices in modeling default, including the standards in the Basel Capital Accord, and therefore may be of more interest to international supervisors. We also describe results of longer-term (two- and three-year) models below in our robustness checks.

[10] Defining failures based on FSCS referrals helps overcome potential endogeneity problems that can arise when failure events are proxied by financial measures, such as a fall in capital ratios or the incurrence of a material loss. Section 3 provides more background on the FSCS referral process and its relationship with credit union insolvency.

results highlight the significance of other sources of risk in addition to capital adequacy that should be considered when developing early-warning criteria for credit union soundness.

The remainder of this paper proceeds as follows. Section 2 provides basic background on the credit union sector in the United Kingdom. Section 3 describes our data set, empirical approach, and framework for evaluating model performance. Section 4 discusses results, while section 5 reviews model performance. Section 6 concludes.

2. Background on U.K. Credit Unions

Credit unions are not-for-profit financial cooperatives, established to meet the economic and social goals of their members. Due to charter restrictions, they do not conduct business with the general public, but instead serve a group of people characterized by a common bond, e.g., belonging to a particular community or sharing the same employer. For these reasons, they are usually concentrated geographically and their members' payment capacities can be subject to local economic conditions.

Credit unions provide savings products and, in general, small consumer-oriented loans. They are subject to stringent lending restrictions on the amounts that they can lend and the periods of such loans, which means that less than 5 percent of all credit unions in the United Kingdom can make residential real estate loans.[11] Importantly, they are also not subject to the same market pressures for growth and earnings performance as for-profit financial institutions. Nevertheless, they still face pressures from regulatory

[11]In particular, small, version 1 credit unions (i.e., those with capital under £1 million and less than 3,000 members), which comprise over 95 percent of the U.K. credit union sector, cannot lend to any one member more than £15,000 in excess of the member's deposit and cannot lend for a period of more than five years where unsecured and ten years where secured. All other credit unions, known as version 2 credit unions, cannot lend for a period of more than ten years where unsecured and twenty-five years where secured. In addition, these institutions cannot have exposure to any one member of more than £15,000 in excess of that member's deposit or an amount equal to 1.5 percent of capital in excess of that member's shares, whichever is greater. These limitations mean that only version 2 credit unions have the ability to make residential mortgage loans.

requirements to maintain minimum levels of capital and liquidity.[12] If a credit union breaches such minimums, it is required to undertake corrective actions or can be subject to more supervisory intervention, including closure.

3. Data and Empirical Approach

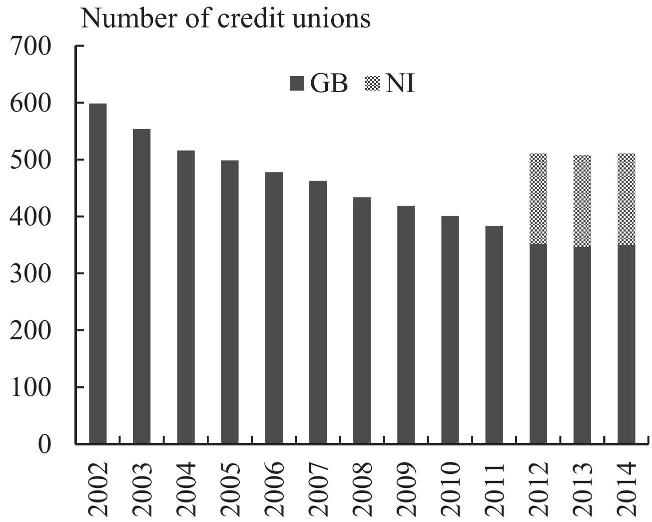
We use annual data (spanning 2002 to 2015) from credit unions' regulatory submissions to the U.K. Financial Services Authority (FSA) (from 2002 to 2013) and Prudential Regulation Authority (PRA) (since 2013). These submissions include details on a credit union's assets and liabilities, profit and loss, solvency, and liquidity positions. The data set is an unbalanced panel containing roughly 6,600 firm-year observations covering information on all credit unions in Great Britain for the years 2002 to 2014 and Northern Ireland for the years 2012 to 2014.

Figure 1 shows that, from 2002 to 2012, the number of credit unions in Great Britain (GB) declined from almost 600 to less than 400 due to mergers with other credit unions and, to a lesser extent, several failures (defined below). The number of credit unions jumped to more than 500 in 2012 as the scope of the U.K. supervisor's oversight expanded to include institutions in Northern Ireland (NI). Figure 2 shows that with the inclusion of NI credit unions, the assets of U.K.-supervised credit unions increased fivefold from less than £0.5 billion in 2002 to more than £2.5 billion at the end of 2014.

We define credit union failure to be the event when an institution was referred to the FSCS for depositor payout. FSCS protection was first extended to U.K. credit unions in 2002 when regulation of the sector was completely transferred from the Registry of Friendly

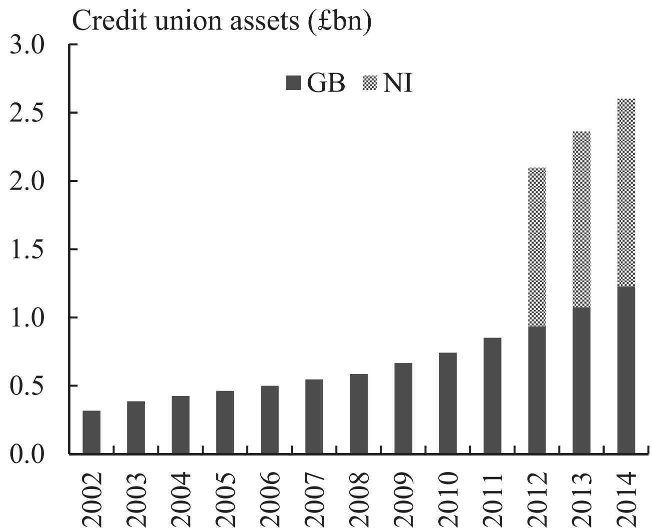
[12] For example, starting in 2013, small, version 1 credit unions (with fewer than 5,000 members and assets under £5 million) were required to maintain capital of at least 3 percent of total assets. This increased to 5 percent for large, version 1 institutions (with members from 5,000 to 10,000 and with assets of between £5 and £10 million). All other (version 2) credit unions were subject to an 8 percent risk-adjusted capital to total asset requirement, where risk-adjusted capital equaled capital plus provision for bad and doubtful debt less the specific provision for bad and doubtful debt. Credit unions were required to hold a liquidity ratio of at least 10 percent at all times.

Figure 1. Number of Credit Unions



Source: Bank of England.

Figure 2. Size of the U.K. Credit Union Sector



Source: Bank of England.

Societies to the FSA. The process of referring a credit union to the FSCS depends on two conditions being met: (i) the credit union must be insolvent (i.e., it reports a capital deficit) and (ii) the credit union has no realistic plan in place to rectify this situation in a timely manner.

Under these circumstances, the only option available is a winding up, which requires an insolvent credit union to appoint an insolvency practitioner to assist with the sale of assets and settle creditors' claims. Once this occurs, regulatory actions are taken to adjust the troubled credit union's permissible deposit-taking and lending activities and to restrict its ability to make any distributions. These measures help prevent capital deficiency from increasing and ultimately protect existing members and the FSCS. Within ten days of appointing an insolvency practitioner, the credit union must register this appointment with the Court, which then issues an official winding-up order, triggering FSCS referral by the PRA (FSA prior to 2013). Because the time between insolvency and referral is short and since the practices have remained unchanged since 2002, we believe that using FSCS referral consistently captures the timing and occurrence of failure relatively well over our sample period. Using this strategy, we identified eighty-five failures, although only sixty-eight are included in our data set due to a lack of corresponding regulatory return information for seventeen institutions. Table 1 reports the annual failure rates from 2003 to 2015, which averaged around 1 percent but varied considerably, peaking in the two years in the immediate aftermath of the crisis.[13]

3.1 Explanatory Variables

Table 2 summarizes our candidate CAMEL variables and the expected association with failure. To gauge capital adequacy, we include a simple capital ratio—equal to total capital as a percentage

[13]We also find that no one region stood out as contributing disproportionately to firm failure. The distribution of failures across the United Kingdom was as follows (with percentage of total in parentheses): Greater London: 5 (7%); East Midlands: 5 (7%); Eastern England: 2 (3%); Northeast: 5 (7%); Northwest: 7 (10%); Northern Ireland: 1 (1%); Scotland: 7 (10%); Southeast: 2 (3%); Southwest: 6 (9%); Wales: 4 (6%); West Midlands: 9 (13%); Yorkshire and the Humber: 6 (9%); not assigned: 9 (13%). Total: 68 (100%).

Table 1. Credit Union Failures (2003 to 2015)

Year	Total Credit Unions	Failures[a]	Failure Rate[b]
2003	625	3	0.48%
2004	584	1	0.17%
2005	543	1	0.18%
2006	532	5	0.94%
2007	512	8	1.56%
2008	497	6	1.21%
2009	468	6	1.28%
2010	452	10	2.21%
2011	428	8	1.87%
2012	405	6	1.48%
2013	518	6	1.16%
2014	514	2	0.39%
2015	514	6	1.17%
Total	6,592	68	1.03%

Sources: Bank of England and authors' calculations.
[a]Institutions referred to the FSCS.
[b]Number of failures divided by total number of credit unions.

of total assets—which is analogous to the leverage ratio for banks. The intuition for its use is that lower capital ratios make institutions more vulnerable to shocks, since they have lower capacity to absorb unexpected loan losses or material declines in asset values. We expect this measure to be negatively associated with the probability of default.

With respect to asset quality, we evaluate four variables. First, we examine the proportion of loans that are in arrears. The arrears rate measures the quality of a firm's existing loan book, with a higher arrears rate indicating potential losses on these loans. We use two distinct ratios: one based on loans between three and twelve months in arrears and another based on loans more than twelve months in arrears. Second, we use the ratio of provisions to loans to capture credit unions' own assessments of losses embedded in their loan portfolios. Third, we consider the ratio of unsecured loans to total assets

Table 2. CAMEL Predictors of Failure

CAMEL Factor	Definition[a]	Expected Association with Failure
Capital Adequacy:		
Simple Capital Ratio	Total Capital/Total Assets	–
Asset Quality:		
Arrears 3–12 Months	Net Liabilities 3–12 Months in Arrears/Total Net Liabilities[b]	+
Arrears > 12 Months	Net Liabilities > 12 Months in Arrears/Total Net Liabilities[b]	+
Provision Coverage	(General + Specific Provisions)/Net Liabilities in Arrears	–
Unsecured Loans	Unsecured Loans/Total Assets	+
Management:		
Size	Natural Log of Total Assets	–
Cost-to-Income	Total Expenditure/Total Income	+
Admin. Expense	Administrative Expense/Total Assets	+
Members	Number of Qualifying Members	+/–
Full-Time Staff	Number of Full-Time Employees	+/–
Earnings:		
Return on Assets	After-Tax Profit (Loss)/Total Assets	–
Liquidity:		
Liquidity Ratio	Total Liquid Assets[c]/Total Relevant Liabilities[d]	+/–
Loans to Deposits	Total Loans/Total Members' Share Balance	+/–

[a]Unless otherwise noted, all measures are expressed as percentages and have been multiplied by 100.
[b]Net liabilities equal total loans plus accrued interest less the members' share balances used to secure the loan.
[c]Total liquid assets equal the sum of qualifying cash and bank balances, investments realizable within eight days, unused committed facilities, and unused overdrafts.
[d]Total relevant liabilities equal the sum of unattached shares and liabilities with an original or remaining maturity of less than three months.

as a proxy for credit risk exposure overall.[14] A higher ratio suggests relatively greater credit risk. We expect these three proxies for asset quality to be positively associated with the likelihood of failure. As a fourth proxy for asset quality, we considered loan loss provisions as a percentage of arrears. The relationship between this measure and failure is ambiguous. On the one hand, a relatively high coverage ratio may imply that credit unions have more than adequately provided for losses embedded in past-due loans. If that is the case, then we would expect the relationship with failure likelihood to be negative. If, on the other hand, higher coverage ratios imply a deterioration in asset quality overall, then we might expect to find a positive association with the likelihood of failure.

We use five different measures to capture the third CAMEL variable, management quality. We use credit union size (measured by the natural log of total assets), and two measures of efficiency, approximated by the ratio of total costs to total revenue (income) and the ratio of total administrative expenses to total assets. We expect size to be inversely correlated with probability of failure based on the idea that larger credit unions may be better diversified across borrowers and geographic location. We expect the two efficiency ratios to be positively associated with failure, with higher (lower) values of these indicators suggesting worse (better) managerial quality. Finally, we include a credit union's number of members and of paid staff as additional proxies for management quality.

Prior research has found that measures of earnings performance are useful in explaining bank failure. To measure earnings performance, we use the ratio of after-tax net income to total assets for our fourth CAMEL proxy. We expect a negative relationship between this ratio and failure.

Finally, for our fifth CAMEL indicator, liquidity, we consider the standard liquidity ratio, defined as liquid assets to total liabilities (see table 2 for definition) and the ratio of loans to deposits. The expected effect of the standard liquidity ratio on failure is, ex ante, ambiguous. On the one hand, it could be inversely associated with the likelihood of failure to the extent that liquid assets provide a useful secondary source of liquidity that credit unions can use to

[14]Regulatory data limitations prevent a finer breakdown of the loan portfolio by type.

Table 3. Summary Statistics[a]

Variables	Obs.	Mean	Std. Dev.	Min.	Max.
Size (Log of Total Assets)	6,601	12.89	1.67	4.46	18.71
Capital Ratio (%)	6,591	8.74	8.17	−11.36	48.01
Arrears 3–12 Months[b] (%)	6,448	9.65	14.33	0.00	97.24
Arrears over 12 Months[b] (%)	6,486	7.49	14.24	0.00	85.71
Unsecured Loans[c] (%)	6,306	41.51	24.81	2.59	100.39
Provision Coverage (%)	5,557	179.16	297.13	36.28	2,200.00
Cost-to-Income Ratio (%)	6,601	82.84	33.73	16.31	230.68
Admin. Expense (%)	6,601	6.82	10.08	0.00	60.43
Membership (000's)	6,578	1.76	13.49	0.00	1,069.79
Full-Time Staff	6,607	2.92	9.87	0.00	663.00
Return on Assets (%)	6,601	1.61	3.47	−13.29	13.27
Liquidity Ratio (%)	6,564	77.07	63.00	6.70	501.99
Loans-to-Deposits Ratio (%)	6,560	68.56	28.35	10.93	159.83
National Unemployment Rate	6,601	6.18	1.27	4.80	8.10
Regional Unemployment Rate	6,018	6.31	1.56	3.40	10.60

[a]Bank-level characteristics based on measures winsorized at the 1st and 99th percentiles.
[b]Percentage of total net loans to members.
[c]Percentage of total assets.

satisfy unexpected liquidity needs. On the other hand, inefficiencies may arise from holding higher proportions of liquid assets, which could weigh on earnings performance and subsequently contribute to the likelihood of failure. Regarding our second measure, the ratio of loans to deposits, while a higher ratio implies a potentially less liquid balance sheet, it could indicate a more efficient use of funding, which could improve earnings and capital. As a result, the expected impact on failure is also, ex ante, ambiguous.

3.2 Descriptive Statistics

Table 3 reports summary statistics on our CAMEL variables, while table 4 reports mean equality tests between failures and survivors from a simple univariate analysis for our main candidate

Table 4. Mean Comparison Tests

Variables	**Failures**		**Survivors**		**Mean Equality Test (t-test, Unequal Variances)**	
	Mean	**Obs.**	**Mean**	**Obs.**	**Difference**	**p-value**
Size (Log of Total Assets)	11.88	421	12.96	6,180	−1.08	0.00
Capital Ratio (%)	4.04	418	9.06	6,173	−5.02	0.00
Arrears 3–12 Months[a] (%)	13.82	463	9.33	5,985	4.49	0.00
Arrears over 12 Months[a] (%)	10.57	466	7.25	6,020	3.32	0.00
Unsecured Loans[b] (%)	11.03	336	7.59	5,401	3.44	0.00
Provision Coverage (%)	50.93	399	40.87	5,907	10.06	0.01
Cost to Income (%)	107.52	421	81.16	6,180	26.36	0.00
Admin. Expense (%)	11.87	421	6.48	6,180	5.39	0.00
Membership (000's)	0.63	419	1.84	6,159	−1.21	0.00
Full-Time Staff	3.22	421	2.90	6,180	0.32	0.83
Return on Assets (%)	−0.55	421	1.76	6,180	−2.31	0.00
Liquid Asset Ratio (%)	74.78	413	77.23	6,151	−2.45	0.32
Loans to Deposits (%)	77.93	416	67.92	6,144	10.01	0.00

[a]Percentage of total net loans to members.
[b]Percentage of total assets.

variables.[15] The tests suggest that failed credit unions are generally smaller, less well capitalized, and less profitable. They also tend to have weaker asset quality (as reflected by higher arrears rates, loan loss provisions, and unsecured loans), poorer management efficiency ratios (i.e., higher cost-to-income ratios), and weaker liquidity positions (higher loan-to-deposit ratios).

3.3 Modeling Failure

Following Shumway (2001), we model failure using a multi-period logit model. Since we are interested in whether the CAMEL variables help anticipate failure regardless of firm or time period, we pool the data across firms and over years. This approach allows for time variation in the explanatory variables and treats a credit union's condition as a function of its latest financial measures (as derived from annual regulatory returns).[16] The probability of credit union failure over the next year is equal to

$$P\left(Y_{i,t}=1|X_{i,j,t-1}\right)=\frac{1}{1+exp(-\beta_0+\sum_{j=1}^{J}\beta_j X_{i,j,t-1})}, \qquad (1)$$

where $Y_{i,t}$ is an indicator variable equal to one if credit union i fails (i.e., referred to the FSCS for payout) in year t and equal to zero if the credit union remains active. The term in the denominator on the right-hand side, $\sum_{j=1}^{J}\beta_j X_{i,j,t-1}$, represents a linear combination of our j explanatory variables, which, discussed below, includes CAMEL factors lagged by one year. For early-warning use, this specification helps to understand what factors differentiate failed from successful credit unions and how their state changes in the run-up to FSCS referral.[17]

---

[15] We winsorized all bank-level variables at the 1st and 99th percentiles to mitigate the influence of extreme outliers.

[16] Poghosyan and Cihak (2011) employ a similar approach to examine the determinants of bank distress in Europe, while Cole and Wu (2009) extend this approach to investigate factors explaining U.S. commercial bank failures during the banking crisis of the late 1980s and early 1990s.

[17] Our baseline model employs explanatory variables lagged one year. As a robustness check (discussed below), we also tested specifications with two- and three-year lags. The tests produced results that were broadly similar but weaker in terms of statistical significance, indicating that the predictive power of these models diminishes as the horizon over which failures are predicted increases.

Table 5. Classification Matrix

Signal Issued One Year Prior (at time t–1)	Forecasted Event $F_{i,t-1}$	Actual Failure Event $Y_{i,t}$	
		Failure	**Non-failure**
Yes ($F_{i,t-1} = 1$)	Failure	True Positive (TP)	False Positive (FP)
No ($F_{i,t-1} = 0$)	Non-failure	False Negative (FN)	True Negative (TN)

To estimate this equation, we consider the possibility that individual credit union observations may be correlated across time and that the errors across institutions may not be identically distributed. Ignoring this possibility would lead to downward-biased estimates of standard errors of the coefficients. To deal with this issue, we use logistic models that are robust to clustering of errors at the firm level.

3.4 Evaluating Model Performance for Supervisory Use

To evaluate our failure models and the potential tradeoffs in the context of early-warning systems, we rely on the standard type I versus type II error tradeoff approach used in the banking literature (e.g., Poghosyan and Cihak 2011; Aikman et al. 2014; Betz et al. 2014). This involved choosing a threshold probability, $\pi \in [0, 1]$, above which our model issues a "signal," warning that a credit union is vulnerable and at risk of failure. To facilitate evaluation of such signals, we transform the probability of failure that derives from our logistic model based on data reported at period $t-1, p_{i,t-1}$, into another binary variable $F_{i,t-1}$ that equals one if $p_{i,t-1}$ exceeds π and zero otherwise. The association between the forecast (signal) $F_{i,t-1}$ and the actual failure, as represented by $Y_{i,t}$, can be summarized using a classification matrix as set out in table 5.

In our case, a type I error occurs when our model fails to classify a credit union failure correctly, i.e., the model does not issue a warning signal when a failure is imminent. A type II error results when a healthy credit union is mistakenly forecast to fail. Relating this to the classification matrix above, type I errors are calculated as $T_1 = FN/(TP+FN)$, and type II errors as $T_2 = FP/(FP+TN)$. To compare and contrast model performance, we use measures from receiver operating characteristic (ROC) curves and the area under the ROC curve (AUROC). The ROC curve plots, for the complete range of threshold probabilities $\pi \in [0, 1]$, the conditional probability of positives to the conditional probability of negatives:

$$ROC = \frac{PR\,(Y_{i,t} = 1|F_{i,t-1} = 1)}{1 - PR\,(Y_{i,t} = 0|F_{i,t-1} = 0)}, \tag{2}$$

where $Y_{i,t}$ and $F_{i,t-1}$ are as defined above. In this regard, the ROC curve shows the tradeoff between the benefits (i.e., avoiding the costs that derive from missing a failed credit union) and the costs from misclassifying too many healthy firms: the higher the AUROC, the better the model.

4. Results

To establish our baseline results, we ran a series of specifications involving CAMEL variables. Table 6 presents the results of models employing CAMEL components separately (models 1–5) and our preferred baseline model (model 6) including measures capturing all five CAMEL factors. The results of the component models are consistent with expectations and generally show that each aspect, when included separately, is statistically significant in explaining failure. Asset size (model 1), capitalization (model 2), and earnings (model 4) are negatively associated with the probability of failure, which accords with findings from bank failure research (e.g., Cole, Cornyn, and Gunther 1995; Cole and Wu 2009; Poghosyan and Cihak 2011).[18] The positive signs on the arrears rates and the proportion of unsecured loans (model 3) imply that the likelihood of failure increases as asset quality deteriorates. The positive sign on the loan-to-deposit ratio (model 5) suggests that the likelihood of failure increases as liquidity risk increases.

[18] We also found that the cost-to-income ratio was positively associated with the probability of failure, suggesting that credit unions with less-efficient management teams (as measured by a higher cost-to-income ratio) are more likely to fail sometime in the upcoming year. We excluded this variable from our baseline specification due to the high correlation (see table 9 in the appendix) with our earnings measure, ROA, to mitigate issues with multicollinearity. This was also true with respect to the provision coverage ratio, which is highly correlated with the arrears measures included in the preferred specification.

Table 6. One-Year Failure Model

Variables	**Model 1 Management**	**Model 2 Capital**	**Model 3 Asset Quality**	**Model 4 Earnings**	**Model 5 Liquidity**	**Model 6 Baseline**
Size	−0.2923***					−0.3027**
	(0.0728)					(0.1535)
Capital Ratio		−0.1591***				−0.0948**
		(0.0366)				(0.0390)
Arrears $>$ 12 Months			0.0242***			0.0141**
			(0.0046)			(0.0065)
Unsecured Loans			0.0264***			0.0218***
			(0.0058)			(0.0072)
ROA				−0.2095***		−0.0682
				(0.0226)		(0.0417)
Liquidity Ratio					−0.0004	0.0035
					(0.0041)	(0.0032)
Loan-to-Deposit Ratio					0.0182***	0.0106*
					(0.0035)	(0.0058)
Constant	−1.0553	−3.9365***	−6.4041***	−4.7022***	−6.1056***	−2.8441*
	(0.8980)	(0.1591)	−0.3741	(0.1509)	−0.4026	(1.6626)

(*continued*)

Table 6. (Continued)

Variables	**Model 1 Management**	**Model 2 Capital**	**Model 3 Asset Quality**	**Model 4 Earnings**	**Model 5 Liquidity**	**Model 6 Baseline**
No. of Observations	5,725	5,720	5,331	5,725	5,666	5,309
Wald chi^2	16.134	18.893	59.536	85.818	27.185	132.137
Probability > chi^2	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
Pseudo R^2	0.0173	0.0922	0.0685	0.0913	0.0257	0.1778
Log Likelihood	−286.489	−256.017	−237.438	−264.918	−278.919	−201.512
AIC	576.979	516.034	480.875	533.835	563.837	419.024
BIC	590.284	529.337	500.619	547.140	583.764	471.641
AUROC Curve	0.6304	0.7698	0.7494	0.7749	0.6760	0.8364

Notes: This table reports results of the logistic regression model $\log[p_{i,t}/(1-p_{i,t})] = \beta_0 + \Sigma_{j=1}^{N}\beta_j X_{j,i,t-1} + \varepsilon_{i,t}$, where $p_{i,t} = Prob(Y_{i,t} = 1|X_{j,i,t-1})$ is the probability that credit union i is referred to the FSCS for payout in year t given the vector of j explanatory variables at time $t-1$. Definitions of all explanatory variables are listed in table 2. Standard errors are reported in parentheses below their coefficient estimates and adjusted for both heteroskedasticity and within correlation clustered at the credit union level. The area under the ROC curve is a measure of how well each specification can distinguish between failures and survivors, with larger areas representing better performance. *, **, and *** indicate significance at the 0.10, 0.05, and 0.01 level, respectively.

The baseline model (model 6) includes all five CAMEL components and shows that the signs on the coefficients are in line with expectations. It also shows that including all CAMEL elements improves the fit and in-sample classification power of the model. Our pseudo R-squared for the baseline model increases from, on average, 0.06 (across models 1–5) to 0.18, which is comparable with the fit of similar models in the bank failure literature (e.g., Cole and White 2012; De Young and Torna 2012; Antoniades 2017). The area under the ROC curve shows that the classification of failures and survivors improves when using all five versus employing just one of the CAMEL aspects. Overall, the results suggest CAMEL factors together are useful in characterizing future failures.

Previous literature shows that macroeconomic conditions may be important drivers of banking crises (e.g., Beck, Demirguc-Kunt, and Levine 2006; Cihak and Schaeck 2007) and individual bank failures (e.g., De Young and Torna 2012). Lower rates of unemployment, for example, can be associated with a less volatile macroeconomic environment and, in turn, a lower likelihood of bank-level failure. To test if macroeconomic conditions affect the viability of credit unions, we include national and regional unemployment rates as additional controls. Table 7 reports the results of including these additional controls. Model 6A shows that the lagged national unemployment rate is positively correlated with failure. Indeed, the inclusion of national unemployment improves the in-sample fit of the baseline model as suggested by increases in the pseudo R-squared measures and the area under the ROC curve for each model.

Credit union failure may also be sensitive to regional indicators, given the lack of diversification of its membership. Model 6B shows that regional unemployment is positively associated with the likelihood of failure. The results also indicate that the fit of the baseline models improves with the inclusion of regional unemployment.[19]

[19]Comparisons of Akaike information criterion (AIC) and Bayesian information criterion (BIC) measures across each of the models also confirm that the inclusion of unemployment measures is especially helpful in improving in-sample fit. Both measures for all models that include unemployment rates are lower than those for the corresponding baseline model.

Table 7. One-Year Baseline Failure Model Augmented with National and Regional Unemployment Rates

Variables	Model 6 One-Year Baseline	Model 6A National Unemployment	Model 6B Regional Unemployment
Size	−0.3027**	−0.5340***	−0.3764**
	(0.1535)	(0.1567)	(0.1522)
Capital Ratio	−0.0948**	−0.1073***	−0.1197***
	(0.0390)	(0.0352)	(0.0389)
Arrears > 12 Months	0.0141**	0.0099	0.0101
	(0.0065)	(0.0064)	(0.0072)
Unsecured Loans	0.0218***	0.0245***	0.0167**
	(0.0072)	(0.0077)	(0.0084)
ROA	−0.0682	−0.0483	−0.0489
	(0.0417)	(0.0411)	(0.0427)
Liquidity Ratio	0.0035	0.0040	0.0054*
	(0.0032)	(0.0030)	(0.0028)
Loan-to-Deposit Ratio	0.0106*	0.0122**	0.0133**
	(0.0058)	(0.0054)	(0.0056)
National Unemployment		0.5333***	
		(0.1309)	
Regional Unemployment			0.4221***
			(0.1179)
Constant	−2.8441*	−3.5394**	−4.7735***
	(1.6626)	(1.7054)	(1.7469)
No. of Observations	5,309	5,309	4,867
Wald chi^2	132.136	152.887	117.1173
Probability > chi^2	0.0000	0.0000	0.0000
Pseudo R^2	0.1778	0.2084	0.1838
Log Likelihood	−201.512	−194.010	−153.162
AIC	419.023	406.021	324.324
BIC	471.641	465.216	382.736
AUROC Curve	0.8364	0.8693	0.8713

Notes: This table reports results of the logistic regression model $\log[p_{i,t}/(1-p_{i,t})] = \beta_0 + \Sigma_{j=1}^{N}\beta_j X_{j,i,t-1} + \varepsilon_{i,t}$, where $p_{i,t} = Prob(Y_{i,t} = 1|X_{j,i,t-k})$ is the probability that credit union i is referred to the FSCS for payout in year t given the vector of j explanatory variables at $t-1$. Definitions of all explanatory variables are listed in table 2. Standard errors are reported in parentheses below their coefficient estimates and adjusted for both heteroskedasticity and within correlation clustered at the credit union level. The area under the ROC curve is a measure of how well each specification can distinguish between failures and survivors, with larger areas representing better performance. *, **, and *** indicate significance at the 0.10, 0.05, and 0.01 level, respectively.

4.1 Robustness Checks

To assess the reliability of our baseline model, we employed a different measure of failure and a different lag structure on our explanatory variables. In general, our baseline results are robust to these changes.[20]

While using the administrative measure related to FSCS referral to define failure can help mitigate endogeneity issues that can arise when relying on financial metrics to proxy failure, it is also possible that our definition of failure could be biased if supervisors take actions ahead of FSCS referral in an attempt to avoid referral and soften losses to the FSCS. To address this possible bias, we employed a different measure of failure: we consider a credit union to have "failed" if it breached a capital-to-asset ratio threshold of 5 percent, 3 percent, or 0 percent.[21] Table 8, columns 1–3, reports the results of estimating separate models for each progressively lower threshold measure. In general, the results are qualitatively similar to the baseline results reported in table 6. However, one finding that stands out in these capital breach models is that the statistical significance of our earnings measure (ROA) is more prominent. This is not surprising given the role that earnings plays in influencing capital directly.

[20]We undertook a number of additional checks but do not report them to save space. Overall, the additional tests show that our baseline models are robust to including both time and region effects; including other macroeconomic controls (annual rates of GDP growth and inflation); and excluding version 2 credit unions, which are subject to more stringent capital requirements and intensive supervisory scrutiny and have the ability to underwrite residential real estate mortgages. All results are available upon request.

[21]Using this definition, our sample includes significantly higher failure rates, on average, compared with the 1 percent under the FSCS definition: 30 percent based on the 5 percent threshold, 17 percent based on the 3 percent threshold, and 5 percent based on the 0 percent threshold. It should be noted, however, that for twelve of the fourteen years in our estimation period, credit unions in the United Kingdom were not subject to a formal capital requirement. Instead, they were only required to operate with a positive net capital position. Beginning in 2013, U.K. credit unions were subject to a 3 percent capital-to-asset requirement, and it is for this reason that we elected to focus on thresholds around this level.

Table 8. Robustness Checks (one-year model with different left-hand-side variable; baseline two- and three-year models)

Variables	5% Threshold (1)	3% Threshold (2)	0% Threshold (3)	Two-Year Baseline (4)	Three-Year Baseline (5)
Size	−0.2499***	−0.2526***	−0.4107***	−0.5101***	−0.5231***
	(0.0320)	(0.0411)	(0.0714)	(0.1442)	(0.1258)
Capital Ratio	−0.4546***	−0.4037***	−0.2772***	−0.0425	−0.0743**
	(0.0291)	(0.0300)	(0.0221)	(0.0328)	(0.0298)
Arrears > 12 Months	−0.0061**	−0.0017	0.0065	0.0088*	0.0160***
	(0.0030)	(0.0035)	(0.0050)	(0.0052)	(0.0054)
Unsecured Loans	0.0009	0.0029	0.002	0.0133	0.0169**
	(0.0023)	(0.0028)	(0.0043)	(0.0083)	(0.0074)
ROA	−0.0504**	−0.0804***	−0.0815***	−0.1010**	−0.0298
	(0.0245)	(0.0262)	(0.0276)	(0.0442)	(0.0419)
Liquidity Ratio	−0.0006	−0.0010	−0.0029	−0.0122**	0.0029
	(0.0012)	(0.0014)	(0.0022)	(0.0054)	(0.0025)
Loan-to-Deposit Ratio	0.0053**	0.0031	0.0172***	0.0081	0.0086
	(0.0023)	(0.0025)	(0.0035)	(0.0079)	(0.0057)
Constant	4.8843***	3.4715***	1.8738**	1.4122	0.7777
	(0.4288)	(0.5147)	(0.8614)	−1.6544	(1.3636)

(continued)

Table 8. (Continued)

Variables	5% Threshold (1)	3% Threshold (2)	0% Threshold (3)	Two-Year Baseline (4)	Three-Year Baseline (5)
No. of Observations	5,309	5,309	5,309	4,556	3,888
Wald chi^2	437.427	322.676	333.4132	119.416	76.308
Probability > chi^2	0.0000	0.0000	0.0000	0.0000	0.0000
Pseudo R^2	0.4122	0.3820	0.4141	0.1571	0.1144
Log Likelihood	−1,909.505	−1,475.452	−585.2786	−220.611	−248.219
AIC	3,835.011	2,966.903	1,186.5572	457.223	512.439
BIC	3,887.628	3,019.520	1,239.1745	508.617	562.565
AUROC Curve	0.9215	0.9162	0.9253	0.8359	0.7891

Notes: Columns 1, 2, and 3 of this table report results of the logistic regression model $\log[p_{i,t}/(1-p_{i,t})] = \beta_0 + \Sigma_{j=1}^{N}\beta_j X_{j,i,t-k} + \varepsilon_{i,t}$, where $p_{i,t} = Prob(Y_{i,t} = 1|X_{j,i,t-k})$ is the probability that credit union i breaches a predetermined capital (to asset) ratio threshold in year t given the vector of j explanatory variables at time $t-1$. Columns 4 and 5 report results of the logistic regression model $\log[p_{i,t}/(1-p_{i,t})] = \beta_0 + \Sigma_{j=1}^{N}\beta_j X_{j,i,t-k} + \varepsilon_{i,t}$, where $p_{i,t} = Prob(Y_{i,t} = 1|X_{j,i,t-k})$ is the probability that credit union i is referred to the FSCS for payout in year t given the vector of j explanatory variables at time $t-2$ and $t-3$, respectively. Definitions of all explanatory variables are listed in table 2. Standard errors are reported in parentheses below their coefficient estimates and adjusted for both heteroskedasticity and within correlation clustered at the credit union level. The area under the ROC curve is a measure of how well each specification can distinguish between failures and survivors, with larger areas representing better performance. *, **, and *** indicate significance at the 0.10, 0.05, and 0.01 level, respectively.

Pragmatic considerations related to supervisory use of early-warning models motivates our focus on a one-year-ahead model. We also evaluate how our baseline results hold under longer, two- and three-year horizons by altering the lag structure on our explanatory variables. Columns 4 and 5 of table 8 report the results of specifications that include variables lagged two and three periods before failure. The signs on all coefficient estimates remain unchanged. Comparisons of the pseudo R-squared and AUROC measures show that fit and classificatory power diminish over longer time horizons.

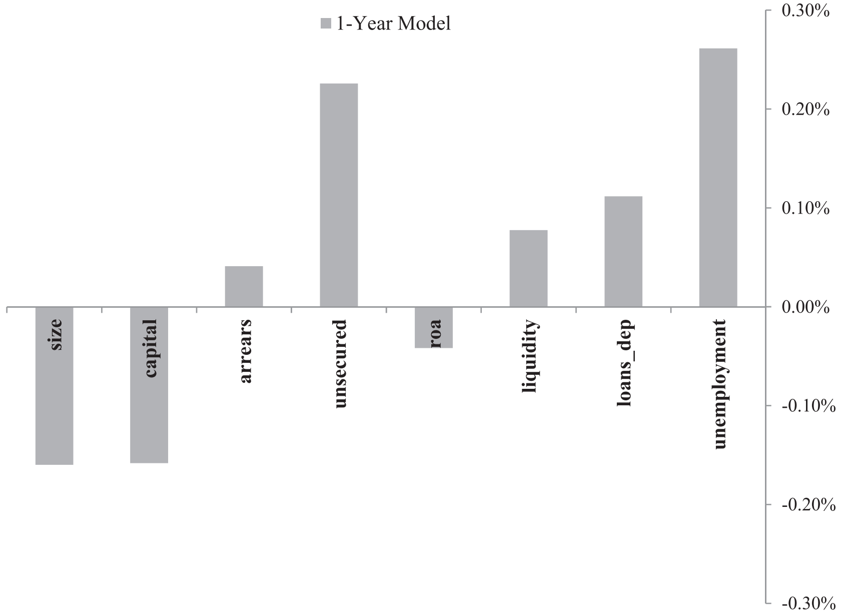
4.2 Economic Significance

The coefficients of the baseline model measure the direction of the impact on the probability of failure. It is difficult, however, to interpret the economic significance of each factor in explaining failure since the magnitude of the impact depends on the initial values of all independent variables and their coefficients. Following standard practice, we derive the economic impact of the individual CAMEL factors by computing the marginal effects at the sample average. In particular, we compute the change in the probability of failure for a one-standard-deviation change in each variable separately for each of the eight variables in the baseline model augmented with national unemployment, holding all other variables constant at their sample average.

Figure 3 shows the relative marginal effects of each covariate in our baseline model, supplemented with national unemployment rates.[22] The shaded bars represent the change in the probability of failure associated with a one-standard-deviation increase in each covariate, holding all other variables at the sample mean. The figure shows that asset quality (proportion of unsecured loans) and national unemployment rates have material influences on the likelihood of credit union failure, while size and capitalization play

[22] For completeness, we have included marginal effects on variables that are not statistically significant: ROA and the liquidity ratio.

Figure 3. Effect of a One-Standard-Deviation Change in Each Variable on Failure Probability

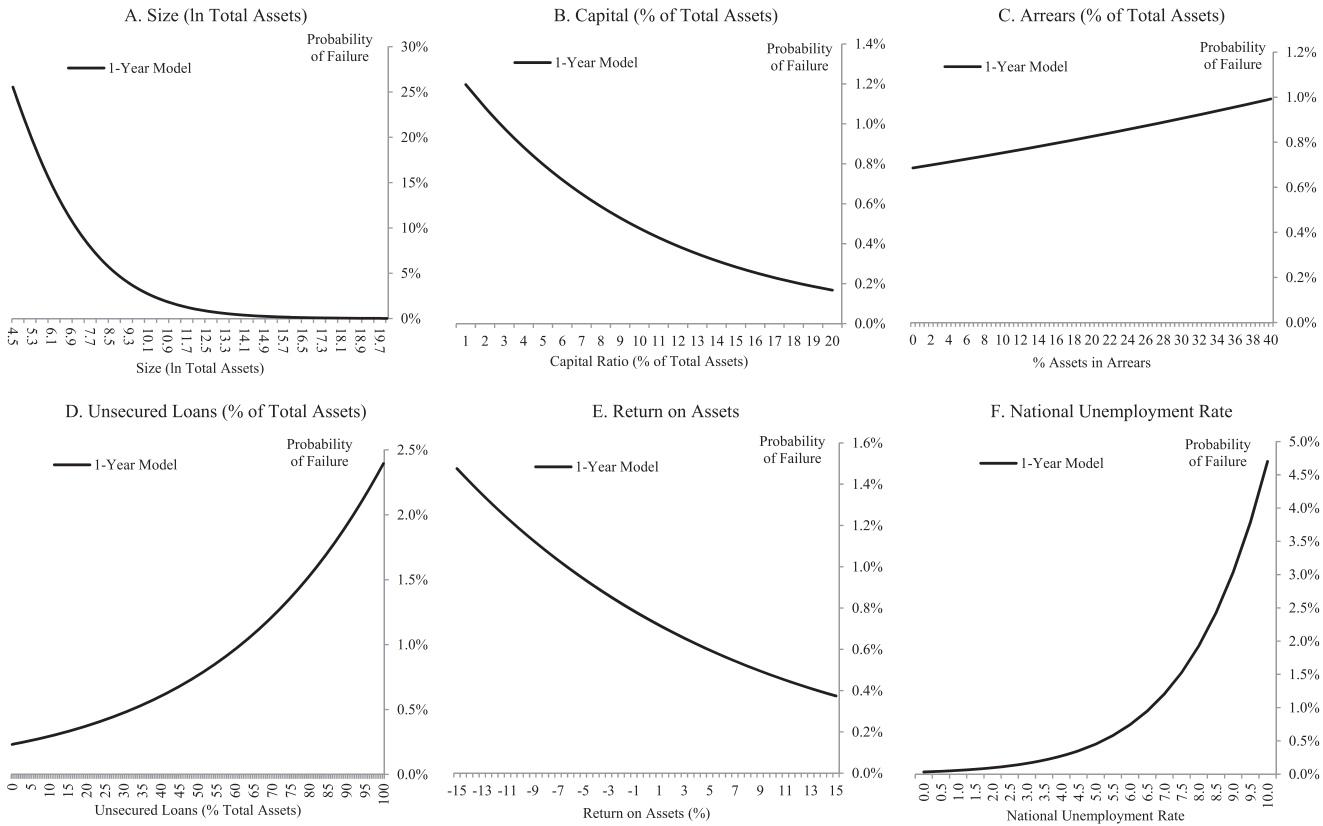


Source: Authors' calculations.

offsetting roles. Liquidity and earnings measures also play minor roles in contributing to failure.

As another way of illustrating economic significance, we report the marginal impacts separately for each of the CAMEL covariates and the national unemployment rate found significant in our augmented baseline models. Figure 4 reports the impacts on the likelihood of failure across a range of plausible values for each covariate separately, while holding all of the remaining independent variables at their sample average. Comparison of marginal effects across the significant determinants of credit union failure suggests that failure probabilities are more responsive to changes in capitalization (panel B) and unsecured loan proportions (panel D). These figures again illustrate the pronounced effects of unemployment and the importance of considering economic conditions for early identification of at-risk credit unions.

Figure 4. Marginal Effects of Explanatory Variables



Source: Authors' calculations.

5. Model Performance

This section reviews the ability of our estimated baseline model to classify failed and surviving credit unions correctly. We focus on evaluating performance in terms of the "usefulness" they provide in terms of minimizing costs associated with type I and type II errors and by comparing the area under receiver operating characteristic (ROC) curves. Better models under this approach would have a higher benefit (i.e., true positive, or "hit rate" on the vertical axis) at the same cost (i.e., false positive, or "false alarm rate" on the horizontal axis). Each false positive along the horizontal axis is associated with a threshold, meaning that the ROC curve measures show performance over all thresholds (not just a predetermined threshold, such as π discussed earlier). The area under the curve measures the likelihood that a randomly chosen failure event is ranked higher than a non-failure. A perfect ranking has an area equal to 1, while random chance has an area under the curve of 0.5.

5.1 In-Sample Classification Performance

The in-sample results reported in table 6 indicate that, while the baseline model has relatively high discriminatory power, performance improves with the addition of unemployment rates. In particular, the area under the ROC curve for the baseline model is roughly 84 percent and increases to 87 percent when augmented with national (or regional) unemployment rates. Overall, the in-sample results suggest that there could be scope for improving model performance by augmenting firm-level characteristics with unemployment rates.

5.2 Out-of-Sample Forecasting Performance

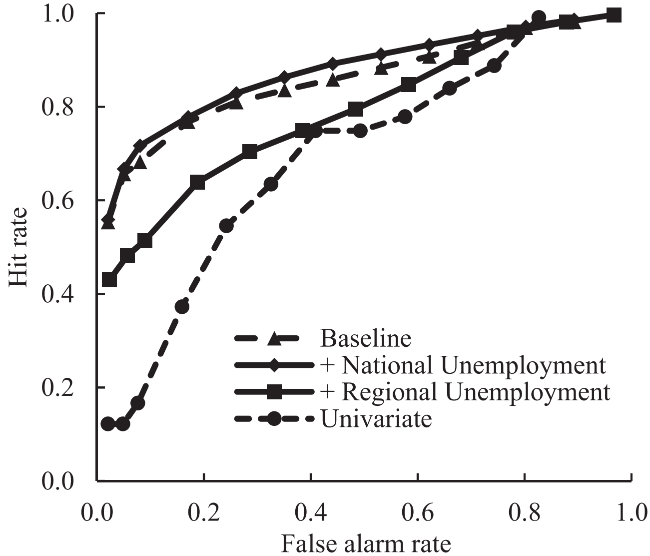
This section evaluates the out-of-sample performance of the models discussed above. More complex models that use more variables to predict failure will by design perform relatively well in in-sample fitting. The more dimensions that are used to explain observed failures, the closer the models will get to replicating these failures. This need not be true out of sample. Indeed, there is a body of evidence that shows that relatively simple models can outperform complex models based on out-of-sample measures (e.g., Haldane and Madouros 2012; Aikman et al. 2014). An aim of this paper is to evaluate the out-of-sample performance of the more complex multivariate models set out above for early-warning use by credit union supervisors.

There is no obvious separate sample on which to evaluate our models discussed above. As such, we rely on data-splitting techniques to separate our single sample of data into two, leaving a "training" sample on which we estimate our model and a "testing" sample on which we evaluate our model's performance in predicting actual credit union failures. This procedure is analogous to how we intend the model to be used in the supervision of credit unions—estimating the model on historical data of past credit union characteristics and failures (analogous to our training sample) and then using the estimated model to rank credit unions, based on their most recent regulatory data, according to their likelihood of failure. All else equal, a model that accurately classifies failure in our testing sample should perform well in predicting the failure of credit unions currently under supervision.

We begin the procedure by randomly selecting a subset of 400 of the roughly 700 different credit unions in our sample. We estimate four separate and distinct models on this training sample: (i) a simple univariate model based on asset size, (ii) our baseline model discussed above that includes only CAMEL covariates, (iii) our baseline model augmented with national unemployment measures, and (iv) our baseline model augmented with regional unemployment measures. We then fix model parameter estimates, and use these to calculate predicted probabilities of default for each of the firms in the testing, or holdout, sample. We define an "at risk" subsample of holdout firms who have a probability of default greater than some predetermined threshold. We use this subsample to calculate "hit rates" and false alarm (type II error) rates for each of the four models. By varying the threshold above which firms are designated at risk, we trace out ROC curves for each of our models. To minimize sampling error, we repeat this process 1,000 times and calculate average ROC curves.

Figure 5 shows out-of-sample testing results for our baseline model. It shows that the univariate model is clearly outperformed by

Figure 5. Out-of-Sample Performance: One-Year Model



Sources: Bank of England and authors' calculations.

each of the multivariate models. The one-year model with regional unemployment performs worse than the other multivariate models, suggesting that for near-term forecasts, regional unemployment may (somewhat surprisingly) degrade model performance and its use for supervisory purposes. The baseline model without regional unemployment performs better but is narrowly outperformed by the model with national unemployment. This is consistent with our in-sample findings discussed earlier where the inclusion of national unemployment is a highly significant factor in characterizing default in the upcoming year.

6. Conclusions

This paper develops an early-warning model for characterizing individual credit union failure in the United Kingdom based on firm-level and macroeconomic indicators of vulnerability. We define failure as the case where a credit institution was referred to the FSCS for depositor payout as part of a formal administrative process. The results show that a small set of firm-level financial CAMEL measures, including a simple non-risk-based capital ratio, asset arrears rates, unsecured loans, return on assets as well as liquid assets, and loan-to-deposit ratios is effective in characterizing potentially troubled credit unions one year in advance of failure. We also find that controlling for regional and national macroeconomic conditions improves in-sample classification and out-of-sample predictive ability.

As credit unions increase in prominence for households that may not have access to traditional forms of finance, understanding what leads to the failure of these institutions will become of increased importance. We believe our paper's results could be of value to supervisors tasked with ensuring the safety and soundness of individual firms and in identifying emerging threats to firm failure. Knowing which firms, and the extent to which the credit union sector overall, exhibit features similar to those that failed previously could help in allocating scarce supervisory resources and in curbing the effects of credit union failures on depositors and the regional economy. This information may also be of interest to the U.K. FSCS and provide benchmark criteria for establishing risk-sensitive levies supporting this compensation scheme.

For macroprudential purposes, the results may give policymakers at least an initial sense of sector resilience. For instance, there may be concerns when the results show a significant proportion of credit unions with high failure probabilities. Aggregating estimated failure probabilities across firms (e.g., on a simple or asset-weighted average basis) and monitoring these over time may also help reveal incipient risks within the sector. Monitoring how the distribution of failure probabilities evolves over time can also shed light on emerging trends and issues that may be of concern to supervisors.

Appendix

Table 9. Correlation Matrix

	Variable	1	2	3	4	5	6	7
1	Size	1.0000						
2	Capital Ratio	0.1370*	1.0000					
3	Arrears 3–12 Months	−0.1014*	−0.0182	1.0000				
4	Arrears > 12 Months	−0.0899*	−0.0554*	0.3376*	1.0000			
5	Provision Coverage	0.0950*	0.0094	−0.2174*	−0.1557*	1.0000		
6	Unsecured Loans	0.2839*	−0.0079	0.0493*	0.0283*	0.0059	1.0000	
7	Admin. Expense	−0.1701*	−0.0429*	0.0490*	−0.0633*	−0.0398*	−0.0396*	1.0000
8	Cost to Income	−0.2404*	−0.2433*	0.1865*	0.1692*	−0.1117*	0.0727*	0.2670*
9	ROA	0.1094*	0.3061*	−0.1536*	−0.1614*	0.0787*	−0.0512*	−0.0981*
10	Loans to Deposits	0.1833*	0.0770*	0.1332*	0.0478*	−0.0045	0.7192*	0.0506*
11	Liquidity Ratio	−0.2743*	0.2685*	−0.0091	0.0616*	−0.0274*	−0.3440*	0.0853*
12	Members	0.0194	0.0694*	0.0251*	0.0201	−0.0278*	0.0453*	0.0882*
13	Full-Time Staff	0.1513*	0.0240*	−0.0251*	−0.0216*	0.0229*	0.0246*	−0.0148

	Variable	8	9	10	11	12	13	
8	Cost to Income	1.0000						
9	ROA	−0.8163*	1.0000					
10	Loans to Deposits	0.1028*	−0.0265*	1.0000				
11	Liquidity Ratio	0.0185	0.0194	−0.2824*	1.0000			
12	Members	0.1117*	−0.1122*	0.0396*	0.0615*	1.0000		
13	Full-Time Staff	−0.0284*	0.0136	0.0266*	−0.0330*	0.0050	1.0000	

Source: Bank of England regulatory data and authors' calculations.
* indicates significance at the 0.10 level.

References

Aikman, D., M. Galesic, G. Gigerenzer, S. Kapadia, K. Katsikopoulos, A. Kothiyal, E. Murphy, and T. Neumann. 2014. "Taking Uncertainty Seriously: Simplicity versus Complexity in Financial Regulation." Financial Stability Paper No. 28, Bank of England.

Antoniades, A. 2017. "Commercial Bank Failures during the Great Recession: The Real (Estate) Story." Working Paper, Department of Finance, NUS Business School, National University of Singapore.

Beck, T., A. Demirguc-Kunt, and R. Levine. 2006. "Bank Concentration, Competition, and Crises: First Results." *Journal of Banking and Finance* 30 (5): 1581–1603.

Betz, F., S. Oprica, T. A. Peltonen, and P. Sarlin. 2014. "Predicting Distress in European Banks." *Journal of Banking and Finance* 45 (August): 225–41.

Cihak, M., and K. Schaeck. 2007. "How Well Do Aggregate Bank Ratios Identify Banking Problems?" IMF Working Paper No. 07/275.

Cole, R. A., B. Cornyn, and J. Gunther. 1995. "FIMS: A New Monitoring System for Banking Institutions." *Federal Reserve Bulletin* 81 (1): 1–15.

Cole, R. A., and J. Gunther. 1995. "Separating the Likelihood and Timing of Bank Failure." *Journal of Banking and Finance* 19 (6): 1073–89.

Cole, R. A., and L. White. 2012. "Déjà Vu All Over Again: The Causes of the U.S. Commercial Bank Failures this Time Around." *Journal of Financial Services Research* 42 (October): 5–29.

Cole, R. A., and Q. Wu. 2009. "Predicting Bank Failures Using a Simple Dynamic Hazard Model." Working Paper.

Demirguc-Kunt, A. 1989. "Deposit-Institution Failures: A Review of Empirical Literature." *Economic Review* (Federal Reserve Bank of Cleveland) 25 (4): 2–18.

De Young, R., and S. Torna. 2012. "Nontraditional Banking Activities and Bank Failures during the Financial Crisis." Research Paper 2012-12, KU Center for Banking Excellence.

Feinberg, R. M. 2003. "The Determinants of Bank Rates in Local Consumer Lending Markets: Comparing Market- and

Institution-Level Results." *Southern Economic Journal* 70 (1): 144–56.

Feinberg, R. M., and D. Meade. 2017. "Economic Benefits of the Credit Union Tax Exemption to Consumers, Businesses, and the US Economy." Report, National Association of Federally-Insured Credit Unions.

Fiordelisi, F., and D. S. Mare. 2013. "Probability of Default and Efficiency in Cooperative Banking." *Journal of International Financial Markets, Institutions and Money* 26 (October): 30–45.

Francis, W. B. 2014. "UK Deposit-taker Responses to the Financial Crisis: What are the Lessons?" Working Paper No. 501, Bank of England.

Haldane, A., and V. Madouros. 2012. "The Dog and the Frisbee." Speech given at the 36th Annual Economic Policy Symposium, sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 31.

Hannan, T. H. 2002. "The Impact of Credit Unions on the Rates Offered for Retail Deposits by Banks and Thrift Institutions." Working Paper, Federal Reserve Board.

Hope, S. 2010. "Promoting the Effectiveness and Efficiency of Credit Unions in the UK." Friends Provident Foundation Report, Roehampton University.

Jones, P. A. 2010. "Stabilising British Credit Unions: A Research Study into the International Rationale and Design of Credit Union Stabilisation Programmes." Research Report, Research Unit for Financial Inclusion, Faculty of Health and Applied Social Sciences, Liverpool John Moores University.

———. 2016. "British Credit Unions: Transformation and Challenges." In *Credit Cooperative Institutions in European Countries: Contributions to Economics,* ed. S. Karofolas, 233–49. Springer International Publishing.

Mare, D. S. 2015. "Contribution of Macroeconomic Factors to the Prediction of Small Bank Failures." *Journal of International Financial Markets, Institutions and Money* 39 (November): 25–39.

McKillop, D., A. M. Ward, and J. O. S. Wilson. 2011. "Credit Unions in Great Britain: Recent Trends and Current Prospects." *Public Money and Management* 31: 35–42.

Poghosyan, T., and M. Cihak. 2011. "Determinants of Bank Distress in Europe: Evidence from a New Data Set." *Journal of Financial Services Research* 40 (3): 163–84.

Sahajwala, R., and P. Van den Bergh. 2000. "Supervisory Risk Assessment and Early Warning Systems." Working Paper No. 4, Basel Committee on Banking Supervision, Bank for International Settlements.

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

Smith, D. M., and S. A. Woodbury. 2010. "Withstanding a Financial Firestorm: Credit Unions versus Banks." Report, Filene Research Institute.

Whalen, G. 1991. "A Proportional Hazards Model of Bank Failure: An Examination of Its Usefulness as an Early Warning Tool." *Economic Review* (Federal Reserve Bank of Cleveland) 27 (1): 21–31.

Wheelock, D., and P. Wilson. 2000. "Why do Banks Disappear? The Determinants of U.S. Bank Failures and Acquisitions." *Review of Economics and Statistics* 82 (1): 127–38.

Wilcox, J. A. 2005. "Credit Union Failures and Insurance Fund Losses: 1971-2004." Economic Letter No. 2005-20, Federal Reserve Bank of San Francisco.

———. 2007. "Determinants of Credit Union and Commercial Bank Failures: Similarities and Differences, 1981-2005." Study, Filene Research Institute.

Revisions to PCE Inflation Measures: Implications for Monetary Policy*

Dean Croushore
University of Richmond
Federal Reserve Bank of Philadelphia (Visiting Scholar)

This paper examines the characteristics of the revisions to the inflation rate as measured by the personal consumption expenditures price index both including and excluding food and energy prices. These data series play a major role in the Federal Reserve's analysis of inflation. We examine the magnitude and patterns of revisions to both PCE inflation rates. We can forecast data revisions in real time from the initial release to the annual revision released in the following year. Policymakers should account for this predictability in setting monetary policy.

JEL Codes: E01, E52, E37.

1. Introduction

In the 2000s, the Federal Reserve (the Fed) changed its main inflation variable from the consumer price index (CPI inflation) to the inflation rate in the personal consumption expenditures price index

*I thank Amanda Smith and Kati Simmons for outstanding research assistance on this project. I also thank participants at the CIRANO workshop on data revisions, the University of Richmond, the Missouri Valley Economics Association, the Federal Reserve Bank of Philadelphia, the American Economic Association, and the Joint Statistical Meetings of the American Statistical Association, as well as Carlo Altavilla, Sharon Kozicki, and Loretta Mester. Thanks to Tom Stark, Mark Watson, Bruce Grimm, and Alan Garner for help with the data. Thanks also to the editor and referees for helpful comments. This paper was written in part while the author was a visiting scholar at the Federal Reserve Bank of Philadelphia. The views expressed in this paper are those of the author and do not necessarily represent the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Please send comments to the author at Robins School of Business, 1 Gateway Road, University of Richmond, VA 23173, or e-mail: dcrousho@richmond.edu.

(overall PCE inflation) and to the PCE inflation rate excluding food and energy prices (core PCE inflation).[1] In 2012, the Fed established a formal inflation target of 2 percent in the overall PCE inflation rate.[2] Unlike the inflation rate based on the consumer price index (CPI), the PCE inflation rate and the core PCE inflation rate are subject to revision, as are all the components of the national income and product accounts. While one might argue in favor of forecasting the CPI inflation rate because it is not revised, the revisions to the PCE inflation rates occur because of additional source data that are better able to determine the nominal level of personal consumption expenditures and how that level is broken down between real consumption and changes in consumer prices.

Monetary policymakers use data on the PCE inflation rate and core PCE inflation rate in making decisions. But those series could be misleading because of large data revisions. For example, consider the core PCE inflation rate as it appeared in May 2002. At the time, inflation (measured as the percentage change in the price level from four quarters earlier) appeared to be falling sharply, as the line labeled “May 2002” in figure 1 shows.

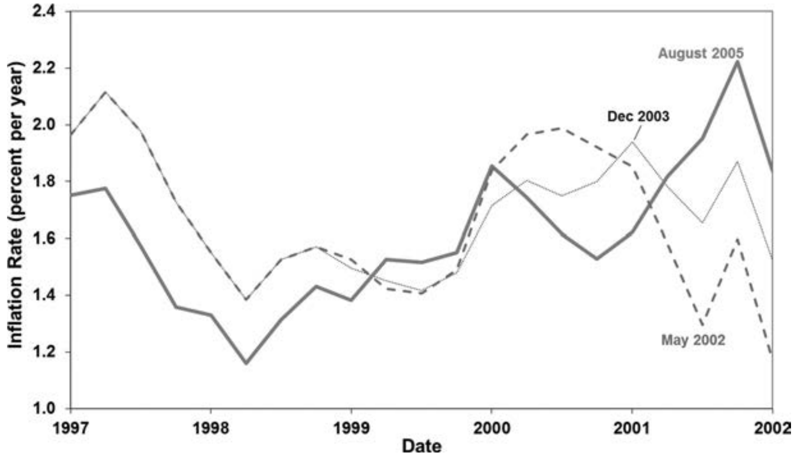
By May 2003, the statement released after the Federal Open Market Committee (FOMC) meeting noted that there could be “an unwelcome substantial fall in inflation.” In a few years, though, the Fed’s worries about the fall in inflation seen in this figure would dissipate because the decline in inflation from 2000 to 2002 would be revised away. For example, in December 2003, the language in the statement after FOMC meetings began to note that the worries about an unwelcome fall in inflation had begun to diminish. As the line labeled “Dec 2003” in figure 1 shows, inflation in 2001 and early 2002 had been revised up by December 2003, so the drop in inflation in early 2002 did not look nearly as worrisome as it had in May 2002.

In fact, a few years later, the worries about a drop in inflation in early 2002 seem misplaced; after the revisions, the data indicated a

[1]The Fed cited three main reasons for the switch: (i) PCE inflation is not subject to as much upward bias as the CPI because of substitution effects; (ii) PCE inflation covers a more comprehensive measure of consumer spending than the CPI; and (iii) PCE inflation is revised over time, allowing for a more consistent time series. See Board of Governors of the Federal Reserve System (2000, p. 4; 2004, p. 3), and Bernanke (2007).

[2]Board of Governors of the Federal Reserve System (2012).

Figure 1. Core PCE Inflation Rate from 1997:Q1 to 2002:Q1: Vintages May 2002, Dec. 2003, and Aug. 2005



rise in inflation from 2000 to late 2001, as the line labeled "August 2005" in figure 1 shows.

Because the PCE inflation rates are revised, as this example illustrates, policymakers need to understand the magnitude of those revisions. This paper seeks to examine those revisions, to determine their overall characteristics, and to investigate the extent to which the revisions might be forecastable. We begin by discussing the data on PCE inflation and its revisions, then analyze a number of tests on the revisions to see if the revisions have desirable characteristics. We use this analysis as a guide to forecasting revisions to PCE inflation in real time. We then discuss the implications of these revisions for monetary policymakers. This paper is the first to develop and use real-time data for core PCE inflation and the first to provide an evaluation of the revision properties of overall PCE inflation and core PCE inflation.

2. Related Literature

Economists have been studying the empirical properties of data revisions since Zellner (1958). Mankiw, Runkle, and Shapiro (1984) found that revisions to the money stock data were reductions of measurement error, so that the initial release of the data was not an optimal forecast of the later revised data. Mankiw and Shapiro (1986) introduced the terminology distinguishing between noise revisions (such as those that occur for the money stock), which are predictable, and news revisions, which are not forecastable. They found that the initial releases of nominal output and real output data are optimal forecasts of the revised data, and thus have news revisions. Mork (1987) suggested that in fact the data released by the government may fit neither the polar case of noise nor the polar case of news, but may be a weighted average of sample information and optimal forecasts. Thus, a test of the correlation of data revisions with information known at the time the data were released provides a general test of well-behavedness of the data. Mork found the initially released data on real GNP growth to be not well behaved, as they are biased downwards and tend to follow their trends more than they should, so that revisions to the data are correlated with existing data known at the time the initial release is produced.

With results like Mork's, which show that revisions are correlated with existing data, it should be possible for the revisions to be predicted in real time. Attempts to forecast such revisions, however, have not always been successful. Much of the time the correlation of revisions with existing data is only known in-sample for a long sample period, but could not be exploited in real time, perhaps because it results from outliers. Faust, Rogers, and Wright (2005) examined data on real output growth for six countries, showing that the revisions are mainly noise. Based on regressions of revised data from initial release to two years later, they were able to predict revisions to the data for most countries. Similarly, Garratt and Vahey (2006) used predictability of U.K. GDP revisions to provide better out-of-sample forecasts of business cycle turning points, using a similar regression approach.

Howrey (1978) showed how to adjust the observation system in a state-space model to account for data revisions. With a similar idea, Conrad and Corrado (1979) used the Kalman filter to form better estimates of revised data on industrial production. Patterson (1995) showed how to exploit the information in past revisions to forecast future revisions using a state-space model. Aruoba (2008) found that most U.S. data revisions are neither pure news nor noise, as suggested by Mork. Aruoba also found that revisions are predictable out of sample, using a state-space model.

How should policymakers respond to data if they know the data may be revised? An early analysis by Aoki (2003) shows that policymakers should be cautious in responding to shocks if the data are measured with error. Orphanides (2001, 2002) argues that simple rules for monetary policy would not have worked as well in real time as they appear to do ex post because of data revisions, and those revisions led to the Great Inflation of the 1970s. More recently, Lubik and Matthes (2016) show that imperfect knowledge about the structure of the economy combined with data misperceptions were key elements that led to the Great Inflation.

Building on the work of Aruoba (2008) and Croushore and Evans (2006), Amir-Ahmadi, Matthes, and Wang (2017) show that data revisions have crucial implications for monetary policy. Their results are similar to those of Croushore and Sill (2016), who show that the responses to macroeconomic shocks are different between initial-release data and final revised data and that the responses to shocks tend to be larger (in absolute value) in final revised data. Thus policymakers considering how to respond to shocks might want to recognize that the impact of the shock might seem small at first but may turn out to be much larger than in the initial estimates of output and inflation.

3. The Data

The real-time data set of the Federal Reserve Bank of Philadelphia, created by Croushore and Stark (2001), is the seminal source for data revisions in U.S. macroeconomic data.[3] Data within any vintage are the exact data available to a policymaker at any given date; generally vintages are based on the data available at mid-month. The data set contains quarterly observations on nominal personal consumption expenditures and real personal consumption expenditures. We use the ratio of these two series to create a real-time data series

[3]See Croushore and Stark (2001) for a description of the overall structure of the real-time data set. Go to the Federal Reserve Bank of Philadelphia's website for the data: http://www.philadelphiafed.org/research-and-data/real-time-center/.

on the overall PCE price index, with vintages available from 1965. The data set did not contain data on the PCE price index excluding food and energy prices, so we collected the data from past issues of the *Survey of Current Business*.[4] The data show the index value of the core PCE price index in each quarter.

From the data on the overall PCE price index and core PCE price index, we create two measures of inflation for each variable, for each observation date and each vintage date, one based on the quarterly inflation rate and a second based on the inflation rate over the preceding four quarters. Our notation for these concepts is $\pi(p, v, t)$ for the overall PCE inflation rate and $\pi^x(p, v, t)$ for the core PCE inflation rate. The first term, p, is the period over which the inflation rate is calculated, with $p = 1$ for quarterly inflation and $p = 4$ for inflation over the preceding four quarters. The second term, v, is the vintage of the data, which is the date on which a policymaker would observe the data; there is a new vintage every month. The third term, t, is the date for which the inflation rate applies. Thus π(4, 2006:M12, 2006:Q3) describes the PCE inflation rate from 2005:Q3 to 2006:Q3, as observed in mid-December 2006, while π^x(1, 2006:M12, 2006:Q3) describes the annualized core PCE inflation rate from 2006:Q2 to 2006:Q3, as observed in mid-December 2006. If P(v, t) describes the level of the overall price index relevant to date t observed in vintage v, then

$$\pi(1, v, t) = \left\{ \left[\left(\frac{P(v,t)}{P(v,t-1)}\right)^4 \right] - 1 \right\} \times 100\% \text{ (the quarterly inflation rate, annualized)},$$

and

$$\pi(4, v, t) = \left\{ \left[\frac{P(v,t)}{P(v,t-4)} \right] - 1 \right\} \times 100\% \text{ (the inflation rate over the previous four quarters)}.$$

[4]These data were produced by this author and research assistants for an early version of this paper and have subsequently been posted as part of the Federal Reserve Bank of Philadelphia's real-time data set.

With these two concepts of PCE inflation and core PCE inflation in hand, we can now describe revisions to the data. Almost always, new data are initially released at the end of January (for the fourth quarter), April (first quarter), July (second quarter), and October (third quarter). The data are revised in each of the following two months after their initial release, then revised in July of each of the subsequent three years, and revised again in benchmark revisions, which occur about every five years. For the first two monthly revisions and the annual revisions, the government agency gains access to additional source data that help produce better values for the data. Benchmark revisions incorporate new data from economic censuses.

Because many revisions occur, we examine a number of different concepts. A variable in the national income and product accounts probably undergoes its greatest revision between its initial release and the first annual revision. That annual revision is the key vintage because the government has access to additional source data that were not available earlier, including income-tax and social-security records and census data that allow for improved sectoral weights, and is thus able to form much more precise measures of income, output, and prices. Another natural revision to consider is that from the initial data release to the latest available series, which for us consists of data from vintage May 2017. In addition, we can consider the data revision from the following year's annual vintage to the latest available data.

Our notation describing the revisions is described as follows. Let $i(1, t) =$ the initial release of $\pi(1, v, t)$ and $i(4, t) =$ the initial release of $\pi(4, v, t)$. Note that these are released at the same time (in the same vintage), but we cannot describe the vintage as "$t+1$" because the vintages are monthly, while the data are quarterly.

Let the annual release of the following year be described as $A(1, t) = \pi(1, v, t)$ and $A(4, t) = \pi(4, v, t)$, where v is the vintage containing the annual revisions in the year after t. Table 1 shows when each annual revision occurs, usually in July of the following year, but sometimes in other months, depending mainly on whether a benchmark revision occurred.

The latest available data come from data vintage May 2017 and are given by $l(1, t) = \pi(1, \text{May 2017}, t)$ and $l(4, t) = \pi(4, \text{May 2017}, t)$.

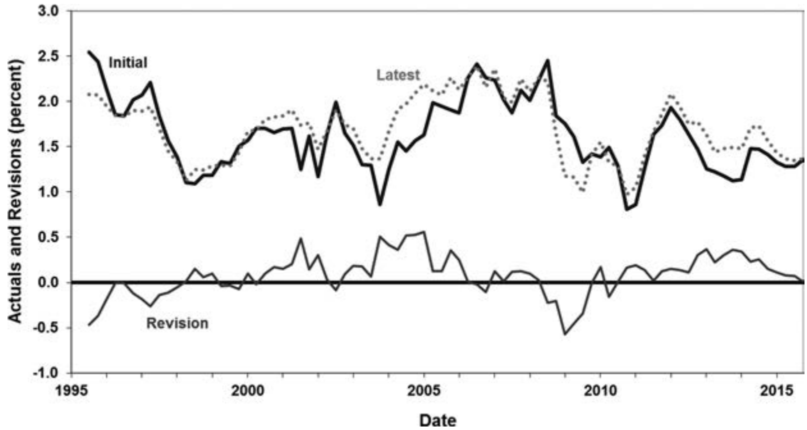
Table 1. Annual NIPA Revision Dates

Year	Annual Revision Date
1965 to 1973	July of following year
1974	January 1976 benchmark revision
1975 to 1978	July of following year
1979	December 1980 benchmark revision
1980 to 1981	July 1982
1982 to 1983	July of following year
1984	December 1985 benchmark revision
1985 to 1989	July of following year
1990	November 1991 benchmark revision
1991	July 1992
1992	August 1993
1993	July 1994
1994	January 1996 benchmark
1995 to 1997	July of following year
1998	October 1999 benchmark revision
1999 to 2001	July of following year
2002	December 2003 benchmark revision
2003 to 2015	July of following year

Given these definitions, the revisions are as follows:

$r(i, A, 1, t) = A(1, t) - i(1, t)$	Revision of quarterly change in PCE inflation from initial release to the first annual revision
$r(i, A, 4, t) = A(4, t) - i(4, t)$	Revision of four-quarter change in PCE inflation from initial release to the first annual revision
$r(i, l, 1, t) = l(1, t) - i(1, t)$	Revision of quarterly change in PCE inflation from initial release to latest available vintage
$r(i, l, 4, t) = l(4, t) - i(4, t)$	Revision of four-quarter change in PCE inflation from initial release to latest available vintage
$r(A, l, 1, t) = l(1, t) - A(1, t)$	Revision of quarterly change in PCE inflation from first annual revision to latest available vintage

Figure 2. Four-Quarter Core PCE Inflation Rate: Initial to Latest Revision and Actuals

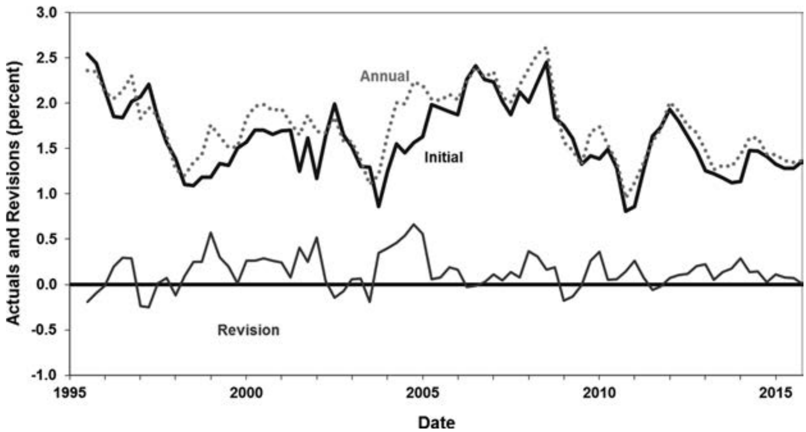


$r(A, l, 4, t) = l(4, t) - A(4, t)$	Revision of four-quarter change in PCE inflation from first annual revision to latest available vintage

We define these revisions for the dates, t, from 1965:Q3 to 2015:Q4, where the latter date is the last one for which an annual revision exists as of this writing. For revisions to core inflation measures, we use the same symbols but add a superscript "x"; for example, $r^x(A, l, 4, t)$ is the revision to the four-quarter core inflation rate at date t between the annual release and the latest available vintage. Revisions to core inflation are available from 1995:Q3 to 2015:Q4 because the core PCE data were first released by the Bureau of Economic Analysis in early 1996.

Figure 2 shows one particularly interesting revision series, which is $r^x(i, l, 4, t)$, the four-quarter revision of core PCE inflation from the initial release to the latest available data in May 2017, shown from 1995:Q3 to 2015:Q4. In the figure, you can see a number of positive revisions to the data from the late 1990s to the late 2000s. Core inflation was revised up on many dates by about 0.5 percentage point, which is a large change, given that the original inflation rate was between 1.0 and 2.5 percent on a number of those dates.

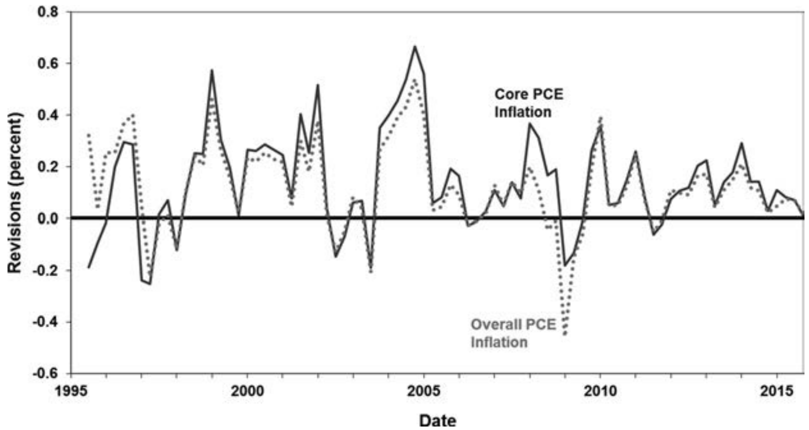
Figure 3. Four-Quarter Core PCE Inflation Rate: Initial to Annual Revision and Actuals



Given that some of the revisions in figure 2 may have arisen from benchmark revisions, which might include changes in definitions, another piece of data to examine is the revision from the initial release to the first annual release. These revisions avoid most of the benchmark changes that might be difficult to anticipate; they are pure revisions caused by additional sample information (except on the few occasions where there was no annual revision because of a benchmark revision). Figure 3 shows the revisions to the core inflation rate from initial to annual release, which is $r^x(i, A, 4, t)$, in our notation. This figure shows that most of the core PCE inflation numbers were revised up from their initial release to the first annual revision. Some revisions are large; for example, in 2004:Q4, the four-quarter core inflation rate was revised up by 0.7 percentage point from its initial release in January 2005 to the annual revision released in July 2006. Overall, figure 3 suggests that an analyst could improve on the initial release of the inflation rate by assuming that it would increase by about 0.12 percentage point by the time of the annual revision, where 0.12 percentage point is the average upward revision from the initial release to the first annual revision.

Finally, we note that revisions to overall PCE inflation and core PCE inflation are broadly similar. For example, figure 4 shows the relationship between revisions from initial release to annual revision,

Figure 4. Revisions to Four-Quarter Inflation Rates: Initial to Annual Revision



for the four-quarter inflation rate in both variables. Both show the same property that the average revision is positive.[5] Taking the revision series as broadly similar for both variables, we take advantage of the fact that the overall PCE inflation series has been available for much longer (since vintage November 1964), to investigate whether problems with the core PCE inflation revisions occur just because the history of the variable is so short. The figure also suggests that the main source of revisions is not revisions to food and energy prices or revisions to the weights in those sectors; otherwise, the lines in figure 4 would look different.

To formalize the ideas suggested by these figures, we proceed to examine the data formally through a variety of statistical tests.

4. Characteristics of the Revisions

Table 2 shows the mean and standard error of the sample mean of each revision concept for one-quarter inflation rates and four-quarter inflation rates, for both core PCE inflation and PCE inflation. We

[5]Forecastable revisions are generally considered to be undesirable by government statistical agencies because they suggest that the methods used to produce the data could be improved.

Table 2. Zero-Mean Test for Revisions

	Core PCE Inflation			Overall PCE Inflation		
Revision	**Mean**	s	p**-value**	**Mean**	s	p**-value**
One-Quarter Inflation Rate						
i_A	0.12	0.037	0.001*	0.10	0.042	0.014*
i_l	0.08	0.044	0.071	0.07	0.058	0.224
A_l	−0.04	0.036	0.248	−0.03	0.054	0.549
Four-Quarter Inflation Rate						
i_A	0.14	0.021		0.10	0.020	
i_l	0.08	0.025		0.04	0.029	
A_l	−0.06	0.019		−0.06	0.029	

Notes: In the table, s is the standard error of the sample mean and the p-value is for the test of the null hypothesis that the mean revision is zero. For core PCE inflation, the sample period is 1995:Q3 to 2015:Q4; for overall PCE inflation, the sample period is 1965:Q3 to 2015:Q4. An asterisk highlights a p-value less than 0.05. Only the one-quarter revision is tested, as the four-quarter revisions are subject to overlapping-observations problems.

expect that the standard error of the sample mean will increase as we look at revisions with more time between measurement dates. So, we expect the standard error to rise as we move from *i_A* to *i_l* (moving down between the first two rows in the table) or as we move from *A_l* to *i_l* (moving up between the third row and second row in the table). Indeed, these patterns do occur, as the table shows.

Figure 4 above suggested that the revision from initial release to annual release in the following year was positive, on average, for core PCE inflation. Formal tests for a zero mean in the revisions to both PCE and core PCE inflation are presented in the columns labeled "p-value" in table 2. The table's results are consistent with what appears to the eye in figures 3 and 4: that the mean revision from initial to annual is significantly above zero. The same is true of the revision from initial to latest available, though the mean revision is not statistically significantly positive. The mean revision from annual to latest available is negative and not statistically significantly different from zero. Thus, it appears that it would be useful

Table 3. Sign Test

Revision	Core PCE Inflation		Overall PCE Inflation	
	s	p-value	s	p-value
i_A	0.68	0.001*	0.59	0.011*
i_l	0.59	0.122	0.55	0.122
A_l	0.49	0.825	0.47	0.398

Notes: s is the proportion of the sample with a positive revision and the p-value is for the test of the null hypothesis that s differs significantly from 0.50 under the binomial distribution. For core PCE inflation, the sample period is 1995:Q3 to 2015:Q4; for overall PCE inflation, the sample period is 1965:Q3 to 2015:Q4. An asterisk highlights a p-value less than 0.05. Only the one-quarter revision is tested, as the four-quarter revisions are subject to overlapping-observations problems.

to forecast a revision to PCE inflation based on a non-zero mean from the initial release to the annual release.[6]

Another useful test for evaluating forecast errors is to examine the signs of forecast errors; we can use the same test to evaluate data revisions.[7] The formal sign test has the null hypothesis that revisions are independent with a zero median, which we test by examining the number of positive revisions relative to the number of observations, assuming a binomial distribution.[8] Results of the sign test applied to various definitions of revisions are shown in table 3. The results are consistent with those in table 2, as non-zero mean revisions are accompanied by non-zero median revisions for the initial to annual revision.

[6]Note that revisions from the initial release to the first annual revision are generally not significant for other macroeconomic variables, such as consumption and output, as I have determined in unpublished research. On the other hand, Aruoba (2008) suggests that some other revisions may be forecastable. For more on the appearance of bias in data revisions, see Croushore (2011).

[7]Research suggests that tests on mean forecast errors may have low power and may be sensitive to outliers, as Campbell and Dufour (1991) suggest, which is why the non-parametric sign test is a useful additional test.

[8]If the revisions were not independent, the sign test would not be appropriate. It is possible that the revisions within one year are not independent. A sign test sampling just once each year shows bias for core PCE inflation in two of four subsamples but no bias for overall PCE inflation in all four quarterly samples. So, the evidence of bias is quite a bit stronger for core PCE inflation.

Table 4. Root Mean Squared Forecast Error for Forecasting Latest Available Actual

Forecast Based On	Core PCE Inflation	Overall PCE Inflation
Initial Release	0.41	0.83
Annual Release	0.33	0.77
DM Test p-value	0.04	0.26

Notes: RMSFE is the root mean squared forecast error from using the vintage concept shown in each row as a forecast of the latest available data. The sample period is 1995:Q3 to 2015:Q4 for core PCE inflation and 1965:Q3 to 2015:Q4 for overall PCE inflation. Only the one-quarter revision is tested, as the four-quarter revisions are subject to overlapping-observations problems. The Diebold-Mariano test reported in the last row is the p-value for a test of the null hypothesis that there is no significant difference in the RMSFE between the two forecasts.

We can think of early releases of the data as forecasts of the latest available data. Assuming a symmetric loss function, a conventional measure of forecastability is the root mean squared forecast error (RMSFE) of each actual series. We should expect that treating the initial release as a forecast of latest-available data will produce a worse forecast (and thus have a higher RMSFE) than using the annual actual as a forecast. Table 4 shows calculations of these RMSFEs for both overall PCE inflation and core PCE inflation. For forecasting the latest available data, the RMSFE declines as we move from initial release to annual release, as expected. The improvement in the RMSFEs is quite large (and statistically significant according to the Diebold-Mariano test) for core PCE inflation, perhaps in part because of the short sample period. The improvement is more modest for overall PCE inflation, and is not statistically significant. So, for core PCE inflation, information in the annual release is valuable, providing a significantly better view of the later measures of the inflation rate than the initial release did. However, for overall inflation (over a longer period), taking the first annual release improves over the initial release as a predictor of the later release, but not significantly so.

Since the seminal papers of Mankiw, Runkle, and Shapiro (1984) and Mankiw and Shapiro (1986), researchers have distinguished between data revisions that are characterized as either containing news or reducing noise. Croushore and Stark (2003) found that many

Table 5. Standard Deviations of Alternative Actuals (in percentage points)

	Core PCE Inflation	Overall PCE Inflation
Initial Release	0.600	2.746
Annual Release	0.575	2.678
Latest Available Release	0.520	2.719

Notes: Each number in the table is the standard deviation of the growth rate of the variable listed at the top of each column for the data set listed in the first column. If revisions contain news, the standard deviation should increase going down a column; if the revisions reduce noise, the standard deviation should decrease going down a column. The sample period is 1995:Q3 to 2015:Q4 for core PCE inflation and 1965:Q3 to 2015:Q4 for overall PCE inflation.

components of GDP were a mix of news and noise. The distinction is an important one, because a variable whose revisions can be characterized as containing news are those for which the government data agency is making an optimal forecast of the future data. Then a revision to the data will not be correlated with the earlier data. On the other hand, a variable that is one for which the revisions reduce noise is one in which the revisions are correlated with the earlier data and are thus forecastable.

News revisions have the property that the standard deviation of the variable over time rises as the variable gets revised; noise revisions have standard deviations that decline as the variable gets revised. So, one simple test of news and noise is to examine the pattern of the standard deviation across degrees to which data have been revised, as we show in table 5. For both core PCE inflation and overall PCE inflation, the standard deviation declines as we move from initial release to annual release, which suggests that revisions from initial to annual are characterized as reducing noise. However, when we move from the annual release to the latest available release, for core PCE inflation the standard deviation declines, while for overall PCE inflation the standard deviation rises, so there is not a consistent pattern.

A second test of news or noise comes from an examination of the correlation between revisions to data and the different concepts of

Table 6. Correlations of Revisions with Actuals

Revision/ Actual	Core PCE Inflation: Initial	Core PCE Inflation: Annual	Core PCE Inflation: Latest Available	Overall PCE Inflation: Initial	Overall PCE Inflation: Annual	Overall PCE Inflation: Latest Available
i_A	$-0.35^{\dagger}$	0.22*	0.08	$-0.22^{\dagger}$	0.01	-0.09
	(4.6)[4.9]	(2.7)[1.8]	(1.0)[0.9]	(3.0)[2.6]	(0.1)[0.1]	(1.1)[1.1]
i_l	$-0.52^{\dagger}$	-0.18	0.17	$-0.18^{\dagger}$	-0.09	0.12*
	(5.5)[5.3]	(1.4)[1.4]	(1.6)[1.8]	(2.9)[2.0]	(1.4)[1.2]	(1.9)[2.1]
A_l	$-0.28^{\dagger}$	$-0.44^{\dagger}$	0.14	-0.03	-0.09	0.19*
	(2.8)[2.6]	(4.6)[3.2]	(1.3)[1.7]	(0.3)[0.3]	(1.1)[1.0]	(2.4)[2.4]

Notes: Each entry in the table reports the correlation of the variable from the data set shown at the top of the column to the revision shown in the first column, with the absolute value of the HAC-adjusted t-statistics in parentheses below each correlation coefficient; the first one in parentheses comes from a regression of the actual value shown in the column header on the revision shown in the first column, while the second one in brackets comes from a regression of the revision on the actual. The sample period is 1995:Q3 to 2015:Q4 for core PCE inflation and 1965:Q3 to 2015:Q4 for overall PCE inflation.
*There is a significant (at the 5 percent level) correlation between the revision and the later data, implying "news."
†There is a significant (at the 5 percent level) correlation between the revision and the earlier data, implying "noise."

actuals. If there is a significant correlation between a revision and later data, then the revisions fit the concept of adding news. But if there is a significant correlation between a revision and earlier data, then the revisions fit the concept of reducing noise. Correlations for core PCE inflation and overall PCE inflation are shown in table 6. For both inflation measures, the results are consistent with the results from table 5 that suggest that the revision from initial to annual reduces noise. Table 5 showed that the standard deviation of the initial release was higher than the standard deviation of the annual release, while table 6 shows that the initial release is significantly correlated with later revisions. Both of those results suggest that the revisions reduce noise. For other concepts, the results are mixed. Most of the remaining cells in table 6 show no significant correlations. For core inflation, the latest available release is not correlated with any of the earlier revisions. For core inflation, the annual release is significantly correlated with the revision from initial to annual, which suggests that the revision represents news, but the annual release is also significantly correlated with the annual to latest release, which suggests noise, and the standard deviation declines between the annual and latest available releases, suggesting noise. For PCE inflation, the annual release is not significantly correlated with any revisions, while the latest available release is significantly correlated with the revision from the annual release to the latest release, suggesting news, which is also supported by the rise in the standard deviation from annual release to latest available release.

Overall, the news-noise results suggest that it may be possible to predict the revisions between the initial release and the annual release for both variables, as those are the revisions for which the evidence is most clear that the revisions reduce noise. Other revisions might be predictable, but the evidence is more mixed.

5. Forecastability of the Revisions

Can we use the results on the characteristics of the revisions to make forecasts of the revisions? One possibility is to use the negative correlation between the initial release of the data and the revision, shown in table 6, to try to forecast the revision. If the correlation is strong enough, we should be able to reduce the root mean squared forecast error of the revision from the values shown in table 4. The problem, of course, is that the correlations shown in table 6 can only be observed after the fact from the complete sample. The question is, could we, in real time, calculate the correlations between earlier release and known revisions and then use that correlation to make a better forecast about the value of the variable after revision?

Given that the revision from the initial release to the annual revision was characterized as reducing noise, we proceed to forecast the initial to annual revision in the following way. Consider a policymaker in the second quarter of 1985 who has just received the initial release of the PCE inflation rate for 1985:Q1.[9] First, use as

[9]We start with the observation for 1985:Q1 to allow enough observations at the start of the period to prevent having too small of a sample period for the regressions that follow, yet still have enough out-of-sample periods to provide a meaningful test. Also, the small sample size for core PCE inflation makes it impossible to do any useful prediction of revisions, so in this section we focus only on overall PCE inflation.

the dependent variable in a regression all the data on revisions from initial release to the annual release that were released through the current period, which in this case gives a sample from 1965:Q3 to 1983:Q4. Regress these revisions on the initial release for each date and a constant term:

$$r(i, A, 1, t) = \alpha + \beta i(1, t) + \varepsilon(t). \tag{1}$$

Use the estimates of α and β to make a forecast of the annual revision that will occur in 1986:

$$\hat{r}(i, A, 1, 1985Q1) = \hat{\alpha} + \hat{\beta} \cdot i(1, 1985Q1).$$

Repeat this procedure for every new initial release from 1985:Q2 to 2015:Q4. Now, based on this forecast of a revision, formulate a forecast of the value of the annual revision for each date from 1985:Q1 to 2015:Q4, based on the formula

$$\hat{A}(1, t) = i(1, t) + \hat{r}(i, A, 1, t). \tag{2}$$

Finally, we ask the question, is it better to use this estimate of the revision based on equation (2), or would assuming no revision provide a better forecast of the annual release? We examine this by looking at the root mean squared forecast error, taking the actual annual release value as the object to be forecasted, and comparing the forecast of that value given by equation (2) with the forecast of that value assuming that the initial release is an optimal forecast of the annual release. Results of conducting such a forecast-improvement exercise are shown in table 7.

The results show that it is indeed possible to forecast the revision that will occur in the annual release. Regression coefficients of equation (1) show a positive constant term and a negative slope coefficient (across all of the regressions, each with a different real-time data vintage used in the regression). The root mean squared error declines by 6.7 percent, which is indicated by the forecast-improvement exercise ratio of 0.933, which is the ratio of the RMSFE based on equation (2) to the RMSFE based on forecasting no revision from initial release to the annual revision. A Diebold-Mariano test suggests that the forecast improvement is significant at the 10

Table 7. RMSFEs for Forecast-Improvement Exercises

Forecast Annual Release	RMSFE
Forecast Based on Initial Release, Equation (2)	0.428
Assume Expected Revision from Initial = 0	0.458
Forecast Improvement Ratio	0.933
Diebold-Mariano *p*-value	0.09

Notes: The forecast-improvement exercise ratio equals the RMSFE for the attempt to forecast the revision divided by the RMSFE when no revision is forecasted (that is, taking the earlier vintage as the optimal forecast of the later vintage). A forecast-improvement exercise ratio less than one means that the revision is forecastable. The sample period is 1985:Q1 to 2015:Q4. The Diebold-Mariano test examines the significance of the forecast difference assuming, in this case, quadratic loss.

percent level; that is a somewhat weak result, but perhaps not surprising given the real-time properties of the test reported by Clark and McCracken (2009).

Suppose we return to the data reported in figure 1 for 2002:Q1. If we had used the regression method based on equations (1) and (2) in real time, we would forecast that the core inflation rate for 2002:Q1 would be revised from 1.17 percent reported in the May 2002 vintage up to 1.55 percent in the annual revision in 2003, an upward revision of 0.38 percent. In fact, the data were revised up to 1.52 percent in the annual revision in 2003, fairly close to our forecast.

6. Implications for Policymakers

Revisions of the size described in the previous section could have a substantial impact on monetary policy. To illustrate, consider a baseline Taylor rule[10] using the overall PCE price index. The Taylor rule says that policymakers set the federal funds rate according to

$$i_t = 2\% + \pi_t + 0.5\tilde{Y}_t + 0.5(\pi_t - \pi^T), \tag{3}$$

[10] The rule was first described by Taylor (1993). It has been the subject of intense research ever since, as summarized in Orphanides (2010).

where i is the nominal federal funds rate, 2 percent represents the equilibrium real interest rate, π is inflation rate over the past four quarters, $\tilde{Y}$ is the output gap, π^T is the inflation target, and the variables with t subscripts are measured at time t. To make the Taylor rule implementable, we use the output gap as measured by the Federal Reserve Board from 1987 to 2012, along with a 2 percent inflation target in the overall PCE inflation rate.[11] We use our real-time data series to determine the Taylor rule's implied federal funds rate using the initial release of the PCE inflation data. Then, we follow the process of forecasting revisions to the PCE inflation rate as described in the previous sessions, and use those estimated revisions as more accurate measures of the inflation rate in equation (3), to see what the Taylor rule would recommend if the Fed had forecasted revisions to the inflation rate.[12] The differences between the alternative Taylor-rule recommendations are shown in figure 5.

Figure 5 shows that, on average over the period from 1988 to 2012, the federal funds rate recommended by the Taylor rule would have been higher by an average of 0.32 percentage point if the Fed had forecasted revisions to the PCE inflation rate, rather than believing that the initial release of the inflation rate was accurate.[13] This should have led to lower rates of inflation in the entire period. Not only would the average federal funds rate be higher if the Fed had forecasted revisions to inflation, but it would have been higher in every single quarter from 1988 to 2012. This occurs because the estimated bias, as measured in real time, leads policymakers to always

[11] The Federal Reserve Board's output gap series is reported by the Federal Reserve Bank of Philadelphia at https://www.philadelphiafed.org/research-and-data/real-time-center/greenbook-data/gap-and-financial-data-set. Currently, the series is not available after 2012 because the Federal Reserve has a five-year embargo on its output gap series. Note that other versions of the Taylor rule (such as that cited by Bernanke 2015), use a bigger weight on the output gap. The output gap is also subject to very large revisions, as Orphanides (2001) notes.

[12] There is uncertainty about the prediction of the revision, so policymakers could consider that element of uncertainty in setting policy optimally. In this analysis, we assume that they set policy following the rule and not considering the implications of added uncertainty on their policy setting.

[13] In several quarters following the Great Recession, 2010:Q4 and 2011:Q1, the federal funds rate recommended by the Taylor rule is positive using the forecasted revisions to the PCE inflation rate but negative using the Taylor rule based on the initial release of the inflation rate.

Figure 5. Difference in Taylor Rule Using Initial vs. Forecasted Inflation

revise up their estimate of inflation over four quarters relative to the initial estimate.[14]

However, the data also suggest that the revisions have been changing their characteristics over time. A plot of the revision from initial to first annual release shows a tendency for the revisions to be smaller in absolute value since about 1994, as figure 6 suggests.

A plot of the moving ten-year standard deviation of the revisions (figure 7) confirms the idea that in the past fifteen to twenty years, the variability in the revisions has declined.

These graphs suggest that the nature of the revisions has changed somewhat over time. Still, although the mean absolute revision has declined somewhat over time (consistent with the decline in the standard deviation in figure 7), the mean revision remains significant

[14]Note that for some one-quarter-ahead forecasts of inflation, the estimated revision is negative. But because the Taylor rule depends on inflation averaged over four quarters, it turns out that all of those revisions are positive during the period in question.

Figure 6. Revision from Initial to Annual

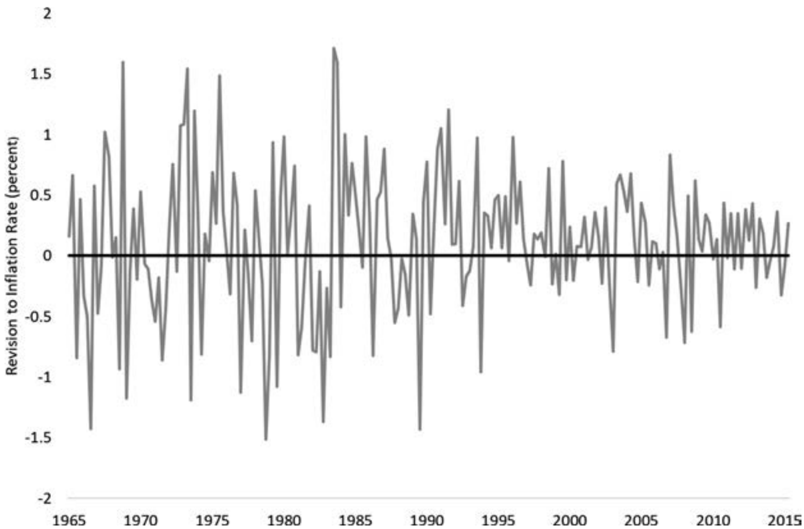


Figure 7. Ten-Year Standard Deviation of Revision from Initial to Annual

and positive, even in recent years, including years since the financial crisis. So, the predictability of revisions remains.

7. Conclusions

We find that, first, the revision from the initial release to the first annual release, for both overall PCE inflation and core PCE inflation, is predictable. Second, the magnitude of those revisions is substantial. In a Taylor rule, the predictability of the revisions would lead to a consistently higher prescribed federal funds rate.

The inflation rate as measured by the percentage change in the PCE price index or the core PCE price index is subject to considerable revisions. Based on this research, monetary policymakers and their staffs may wish to understand the nature of these revisions and factor in the possibility of revisions to the data in making decisions about monetary policy. Clearly, revisions to the PCE inflation rate data are forecastable, and policymakers and economists should forecast revisions to give them a view of what the inflation picture is likely to look like after the data are revised. However, policymakers should also consider uncertainty in the data arising from data revisions, which should cause them to attenuate the strength of their response to new data that are measured with error, as suggested by Sack and Wieland (2000).

References

Amir-Ahmadi, P., C. Matthes, and M.-C. Wang. 2017. "Measurement Errors and Monetary Policy: Then and Now." *Journal of Economic Dynamics and Control* 79 (June): 66–78.

Aoki, K. 2003. "On the Optimal Monetary Policy Response to Noisy Indicators." *Journal of Monetary Economics* 50 (3): 510–23.

Aruoba, S. B. 2008. "Data Revisions Are Not Well Behaved." *Journal of Money, Credit and Banking* 40 (2–3): 319–40.

Bernanke, B. S. 2007. "Federal Reserve Communications." Speech at the Cato Institute 25th Annual Monetary Conference, Washington, D.C., November 14.

———. 2015. "The Taylor Rule: A Benchmark for Monetary Policy?" Brookings Blog, April 28. Available at https://www.

brookings.edu/blog/ben-bernanke/2015/04/28/the-taylor-rule-a-benchmark-for-monetary-policy.

Board of Governors of the Federal Reserve System. 2000. "Monetary Policy Report to the Congress" (February 17). Available at http://www.federalreserve.gov/boarddocs/hh/.

———. 2004. "Monetary Policy Report to the Congress" (February 11). Available at http://www.federalreserve.gov/boarddocs/hh/.

———. 2012. "Federal Reserve Issues FOMC Statement of Longer-Run Goals and Policy Strategy." Press Release, January 25.

Campbell, B., and J.-M. Dufour. 1991. "Over-Rejections in Rational Expectations Models: A Non-Parametric Approach to the Mankiw-Shapiro Problem." *Economics Letters* 35 (3): 285–90.

Clark, T. E., and M. W. McCracken. 2009. "Tests of Equal Predictive Ability with Real-Time Data." *Journal of Business and Economic Statistics* 27 (4): 441–54.

Conrad, W., and C. Corrado. 1979. "Application of the Kalman Filter to Revisions in Monthly Retail Sales Estimates." *Journal of Economic Dynamics and Control* 1 (2): 177–98.

Croushore, D. 2011. "Frontiers of Real-Time Data Analysis." *Journal of Economic Literature* 49 (1): 72–100.

Croushore, D., and C. L. Evans. 2006. "Data Revisions and the Identification of Monetary Policy Shocks." *Journal of Monetary Economics* 53 (6): 1135–60.

Croushore, D., and K. Sill. 2016. "How Data Revisions Respond to Macroeconomic Shocks: Implications from a DSGE Model." Working Paper (May).

Croushore, D., and T. Stark. 2001. "A Real-Time Data Set for Macroeconomists." *Journal of Econometrics* 105 (1): 111–30.

———. 2003. "A Real-Time Data Set for Macroeconomists: Does the Data Vintage Matter?" *Review of Economics and Statistics* 85 (3): 605–17.

Faust, J., J. H. Rogers, and J. H. Wright. 2005. "News and Noise in G-7 GDP Announcements." *Journal of Money, Credit and Banking* 37 (3): 403–19.

Garratt, A., and S. P. Vahey. 2006. "UK Real-Time Macro Data Characteristics." *Economic Journal* 116 (509): F119–F135.

Howrey, E. P. 1978. "The Use of Preliminary Data in Econometric Forecasting." *Review of Economics and Statistics* 60 (2): 193–200.

Lubik, T. A., and C. Matthes. 2016. "Indeterminacy and Learning: An Analysis of Monetary Policy in the Great Inflation." *Journal of Monetary Economics* 82 (September): 85–106.

Mankiw, N. G., D. E. Runkle, and M. D. Shapiro. 1984. "Are Preliminary Announcements of the Money Stock Rational Forecasts?" *Journal of Monetary Economics* 14 (1): 15–27.

Mankiw, N. G., and M. D. Shapiro. 1986. "News or Noise: An Analysis of GNP Revisions." *Survey of Current Business* 66 (5): 20–25.

Mork, K. A. 1987. "Ain't Behavin': Forecast Errors and Measurement Errors in Early GNP Estimates." *Journal of Business and Economic Statistics* 5 (2): 165–75.

Orphanides, A. 2001. "Monetary Policy Rules Based on Real-Time Data." *American Economic Review* 91 (4): 964–85.

———. 2002. "Monetary-Policy Rules and the Great Inflation." *American Economic Review* 92 (2): 115–20.

———. 2010. "Taylor Rules." In *Monetary Economics*, ed. S. N. Durlauf and L. E. Blume. London: Palgrave Macmillan.

Patterson, K. D. 1995. "A State Space Approach to Forecasting the Final Vintage of Revised Data with an Application to the Index of Industrial Production." *Journal of Forecasting* 14 (4): 337–50.

Sack, B., and V. Wieland. 2000. "Interest-Rate Smoothing and Optimal Monetary Policy: A Review of Recent Empirical Evidence." *Journal of Economics and Business* 52 (1–2): 205–28.

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

Zellner, A. 1958. "A Statistical Analysis of Provisional Estimates of Gross National Product and Its Components, of Selected National Income Components, and of Personal Saving." *Journal of the American Statistical Association* 53 (281): 54–65.

Cross-Border Macroprudential Policy Spillovers and Bank Risk-Taking*

Fergal McCann[a] and Conor O'Toole[b,c]
[a]Central Bank of Ireland
[b]Economic and Social Research Institute
[c]Trinity College Dublin

We test whether there was a change in risk-taking by Irish banks in the U.K. mortgage market following the introduction of macroprudential limits on loan-to-value (LTV) and loan-to-income (LTI) ratios in Ireland in early 2015. Using confidential loan-level data on lending in the Irish and U.K. mortgage markets, we provide evidence of risk spillovers whereby Irish banks increased their LTV and LTI ratios on lending abroad in response to the regulatory macroprudential tightening at home. We find heterogeneous effects across groups of borrowers, with LTVs and LTIs increasing most for first-time homebuyers (FTBs). We estimate that, relative to a control group of local lenders, the probability of a high-risk loan being issued by an Irish bank in the U.K. mortgage market increased by 16 percent overall, and by 28 percent in the FTB segment, after the policy introduction.

JEL Code: G21.

*We would like to thank Katharina Bergant, Diana Bonfim, Martin Brown, Steve Cecchetti, Matteo Crosignani, Giovanni Dell'Ariccia, Christina Kinghan, Claire Labonne, Philip Lane, and Steven Ongena for comments and suggestions. We also thank participants at the IMF conference on spillovers from macroeconomic policies, April 2017, the ESRI seminar series, the Bank of England-CEPR conference on Competition and Regulation in Financial Markets, as well as the ECBN conference on macroprudential policy at the Bank of Slovenia, October 2018. The views presented in this paper are those of the authors alone and do not represent the official views of the Economic and Social Research Institute, the Central Bank of Ireland, or the ESCB. Any remaining errors are our own.
Author e-mails: fergal.mccann@centralbank.ie; conor.otoole@esri.ie.

1. Introduction

As the implementation of macroprudential policies, which aim to promote system-wide financial resilience, has become more widespread, policymakers and researchers have paid attention to the possibility that the prudential benefits of regulation of a bank's activity in one sector or country may be eroded through changes in behavior not covered by the regulation. The recent empirical literature has identified a number of mechanisms through which such "spillovers" or "leakages" may occur. Firstly, a domestic regulator may deploy a regulatory instrument, such as those relating to a bank's capital level, which only apply to domestically resident lenders. In such cases, total credit growth in the domestic market may not decrease if unregulated entities such as foreign-owned branches or non-banks "step into the gap" by increasing their lending to offset the credit reduction among regulated firms. Aiyar, Calomiris, and Wieladek (2014) show that precisely this form of prudential spillover arose from bank-varying changes to minimum capital requirements imposed on U.K. lenders in the 1990s and 2000s.

A second form of spillover from prudential regulation may occur when the regulatory instrument deployed targets a specific product or sector within the domestic credit market. A prominent example of such an instrument is the "borrower-based measure," such as a limit on the origination loan-to-value (LTV), loan-to-income (LTI), or debt-service-to-income (DSTI) ratio. Such an instrument targets all loans in the domestic market, regardless of the ownership or residency of the lender. The possible spillover that may arise from such a restriction is the mirror image of the first case: now, banks with operations in multiple countries or asset classes may respond to macroprudential product-based restrictions in the regulated market by changing the volume or profile of their lending in unregulated markets, such as foreign mortgage markets or both domestic and foreign commercial lending markets. Gaining an understanding of the nature and extent of spillovers is an important element of the emerging literature on the effectiveness of macroprudential policy.

In this paper, we use loan-level data for the United Kingdom and Ireland to test for cross-border spillovers from macroprudential regulations in the mortgage market. Using loan-level data on mortgages from Irish-headquartered banks active in both the Republic of Ireland and the United Kingdom, coupled with loan-level data on U.K.-only lenders, we test whether there was a change in risk-taking by Irish banks in the United Kingdom following the introduction of macroprudential regulations in Ireland in February of 2015. Our study is the first to identify this credit risk spillover channel through cross-border lending activity following macroprudential regulations using loan-level information.

Our identification strategy is twofold. First, we use a simple difference-in-difference strategy with loan-level data to compare the lending activity of Irish banks in the United Kingdom with their Irish lending before and after the introduction of macroprudential limits on loan-to-value and loan-to-income ratios in their Irish home market. This comparison provides a first test of whether the behavior of the Irish banks diverged between the two jurisdictions concurrent with the change in regulatory regime in Ireland. Second, to assess whether the risk-taking behavior of the Irish banks changed relative to other banks in the United Kingdom after the policy change, we compare Irish banks' lending conditions (LTV and LTI ratios) in the United Kingdom with that of other mortgage lenders for whom loan-by-loan mortgage data are available from the European Central Bank's European Data Warehouse (EDW). In this second set of specifications, foreign activity of Irish entities is being compared with the activity of local lenders completely unaffected by the Irish mortgage restrictions. We also exploit the granularity in our data by considering whether the Irish banks' lending conditions changed for different buyer types in the market. We focus separately on first-time homebuyers (FTBs), second and subsequent home buyers (SSBs), and refinancing borrowers (Refi).

Two particular econometric issues arise in both our difference-in-difference specifications. First, our objective is to identify supply-side shifts in lending conditions following the regulations. As the loan data are repeated cross-sections of newly issued loans, we are not able to follow the individual borrower over time and include a within-group fixed effect which would directly identify credit demand factors. While such data are more commonplace for enterprises (De Jonghe et al. 2016), it is more difficult to obtain for households, as house purchase decisions are infrequent and lumpy (most households purchase one or two houses in their lifetimes). Indeed, to deal with this issue, the optimal data set would require the universe of both U.K. and Irish micro data in a joint data set with unique U.K. and Irish household identifiers. However, these data sets are not currently compiled by any regulatory authority or other body.

To operationalize the best identification strategy given our data—in particular, dealing with confounding demand-side factors—we saturate the model with a rich set of "loan-borrower type" fixed effects to attempt to purge credit demand effects at as disaggregated a level as possible. This method follows the spirit of Auer and Ongena (2016), De Jonghe et al. (2016), and Khwaja and Mian (2008) and allows us to aim at an identification of changes in the bank risk appetite from a supply side following the policy shift. To create the loan-borrower fixed effects, we use the following borrower and loan characteristics: three borrower age groups for borrowers aged eighteen to thirty-five, thirty-six to forty-five, and forty-six to sixty-nine; year-country house price quartiles; year-country income quartiles; an indicator for four property types (detached houses, semi-detached, terraced houses, apartments), and an indicator for capital cities and their surrounding commuter zones (taking a value of one for Dublin and the mideast regions in the Republic of Ireland (ROI), and London and the (outer) southeast regions in the United Kingdom). This leads to 2,382 groups which are included in the final model once we exclude all cells with one loan.

Second, the repeated cross-sectional nature of the data pose the risk of changes in borrower composition that are affected by the policy. To deal with this issue, we use a multinomial logit inverse probability weighting technique as in Stuart et al. (2014) to deal with selection on observables and to restore the sample composition to be comparable to the pre-policy treatment group.

We find evidence that Irish banks loosened LTVs and LTIs for their U.K. mortgage lending relative to both their Irish lending and the broader lending activity in the U.K. market following the policy change in Ireland. This result demonstrates a risk-shifting channel by which Irish banks take more risk abroad following prudential limits at home. For LTV ratios, we find a 5.3 percentage point difference in the average LTV between Irish banks' ROI and U.K. lending, with the impact strongest for first-time buyers and second and subsequent buyers, relative to buy-to-let and refinance mortgages.

When local U.K. mortgage lenders are used as a control group, the LTV ratios for Irish banks are shown to increase by 1.1 percentage points overall, 3.1 percentage points for first-time buyers (FTBs) and 1.4 percentage points for second-time buyers. There is no difference for refinancing loans. Considering LTIs, we again find evidence of a credit loosening by Irish banks relative to other banks in the United Kingdom after the policy introduction, with this effect limited to borrowers in the FTB segment. Relative to a mean LTI ratio of 2.6 in the United Kingdom over our regression sample period, we estimate a sizable loosening in response to the policy of 0.45.

Our final set of regressions focuses on the probability of issuance of high-risk loans. One could argue that, rather than looking at average LTVs and LTIs, the removal of loans from the right tails of the LTI and LTV distribution may be more relevant when considering the goals of macroprudential policy (to enhance the resilience of borrower and bank balance sheets).[1] We therefore explore whether the Irish banks undertook more "high-risk" lending in the United Kingdom after the Irish macroprudential restrictions were introduced. We define a "high-risk" loan in the United Kingdom to be one with either an LTV above 85 percent or an LTI above 4 (covering between 10 and 15 percent of loans across our sample). We find sizable effects using this approach: Irish banks increased the probability of a high-risk loan by 16 percent relative to the counterfactual, with this effect being 28 percent in the FTB segment. That FTB borrowers experienced the greatest relaxation in credit standards is also noteworthy in terms of understanding the impact of macroprudential regulations, given that FTBs are often faced with downpayment constraints (Engelhardt 1994) and are often the most affected by tightening credit terms. It is clear from this evidence that the spillovers occurred with an increase in credit risk for the Irish lenders, which runs counter to the objectives of the original policies and highlights the difficulties for national regulations in a globalized banking environment.

Our approach is similar in spirit to that of Auer and Ongena (2016), who focus on a cross-asset-class, rather than cross-border,

[1] See Kelly and O'Toole (2018) for discussion.

version of an outward spillover. Using a residential-mortgage-related activation of the countercyclical capital buffer in Switzerland in 2012 as an experiment, they provide evidence that commercial lending increased disproportionately among those lenders with the greatest ex ante mortgage market exposure. Such a finding tallies with the results presented here: where product-specific or asset-class-specific regulation is imposed, lenders may respond by loosening credit conditions on unregulated parts of their balance sheet.

The evidence from the emerging literature on international spillovers from macroprudential regulation is mixed. Buch and Goldberg (2016) report on the results of fifteen separate country studies investigating the international spillovers of prudential policy changes, and find evidence that the effects of prudential instruments sometimes spill over through cross-border bank lending. They show further that bank-specific factors like balance sheet conditions and business models are important drivers of heterogeneity in spillovers, but by no means do they suggest that outward spillovers are a universal consequence of prudential regulation. Using aggregate cross-border lending data, Houston, Lin, and Ma (2012) find that bank flows are positively related to the number of activity restrictions and the stringency of capital regulation imposed on banks in their source country, and negatively related to restrictions and regulations in the recipient country, suggestive of spillovers in activity towards jurisdictions with less regulation. Ongena, Popov, and Udell (2013) show that when more general forms of bank regulation such as barriers to entry and restrictions on bank activity are tighter at home, a multinational bank is more likely to have looser credit standards in its foreign lending, prima facie evidence of an outward risk spillover from domestic regulation.

Certain papers have shown how domestic capital tightening can have contractionary effects on foreign lending of global banks (Aiyar et al. 2014). In this vein, Tripathy (2017) uses municipality-level data in Mexico to test the impact of macroprudential regulation in Spain on the Mexican subsidiaries of Spanish banks. The research finds that capital regulatory tightening in Spain caused a drop in the supply of household credit in Mexico following the regulatory shock. This "outward transmission" leads to less risk-taking abroad as a response to domestic capital regulation, which is in direct contrast to the cross-border "outward spillover" being tested here, where lenders take more risk in foreign markets in response to LTV and LTI restrictions that have effects only in the domestic mortgage market.

The aforementioned studies all test for the presence of spillovers from policy changes on the volume of lending. Our research differs from the extant literature in that it explores the change in risk-taking by explicitly measuring the LTV and LTI ratios of new loans. In this vein our paper is close to Demyanyk and Loutskina (2016), who test whether differences in the regulatory regime between deposit-taking institutions and mortgage companies in the United States resulted in regulatory arbitrage to loosen credit conditions. They find that subsidiary mortgage companies of broader banking groups originate riskier mortgages at looser credit conditions than their deposit-taking arms. The argument is predicated on the fact that banks have a particular risk appetite and, when faced with a regulatory regime which limits risk-taking, they will find ways to circumvent the regulations and continue to take the risk.

In addition to the aforementioned studies that directly address spillovers from domestic financial regulations, our study is also related to that by Reinhardt and Sowerbutts (2015), who show that "inward spillovers" appear to be prevalent after domestic macroprudential regulation, in the sense that foreign bank lending to domestic non-banks increases after a domestic macroprudential capital action. The authors, however, find no evidence that such inward spillovers occur when domestic product regulations such as LTV restrictions are implemented. Our research is also linked to the burgeoning literature on the impact and effectiveness of macroprudential policies (Cerutti, Claessens, and Laeven 2017; Cizel et al. 2016; Vandenbussche, Vogel, and Detragiache 2015) and the literature considering the impact on the mortgage market of bank capital requirements (Basten and Koch 2015; Jimenez et al. 2017).

The paper proceeds as follows: Section 2 introduces the policy and economic context in Ireland and the United Kingdom in the period around the introduction of macroprudential mortgage restrictions by the Central Bank of Ireland; section 3 presents a model of Irish banks' risk-taking at home and abroad; section 4 presents a model comparing ROI banks' risk-taking in the United Kingdom with that of local competitor banks; section 5 assesses the impact of the regulations on the probability of high-risk loans being issued; and section 6 concludes.

2. Policy and Economic Context

2.1 Macroprudential Measures in the Irish Mortgage Market

The role played by mortgage finance in the global financial crisis has been well documented (Adelino, Schoar, and Severino 2016; Duca, Muellbauer, and Murphy 2010, 2011). Loosening credit standards led to a buildup of vulnerabilities and a wave of mortgage defaults in the United States and other markets with housing booms like Ireland and Spain. The outcome was the onset of systemic financial crises in these countries. To prevent a reoccurrence, macroprudential policies on borrower credit such as LTV, LTI, and debt service limits (DSRs) have become more common in the post-crisis era. Such measures have been in place in many Asian economies following the financial crisis in the late 1990s into early 2000s (International Monetary Fund 2013). However, their adoption has become more widespread since the global financial crisis. In 2015, the European Systemic Risk Board (ESRB) reported that a total of seventeen European countries had LTV limits in place, with a number of countries also having income restrictions (ESRB 2016).

On February 9, 2015, the Central Bank of Ireland introduced measures to limit the LTV and LTI ratios applying to new residential mortgage lending in the Irish republic. With many mortgage holders struggling to meet their repayments, and an arrears rate still among the highest in Europe, the measures were enacted with a view to improving the resilience of households and the banking sector to financial shocks and reducing the risk of future bank credit and house price spirals. In Ireland, following the financial crisis, house prices fell by between 50 and 60 percent from peak levels, while 20 percent of mortgages were in arrears at the peak of the crisis in 2013. This highlights the scale of the difficulties in Ireland and the requirement for policies to safeguard financial stability into the future.

Table 1 provides an overview of the measures as they applied to FTBs and second and subsequent buyers (SSBs) from the implementation date until the end of 2016.[2] For FTBs, a sliding LTV

[2]In late 2016, an amendment to the regulations was announced which loosened credit conditions for first-time buyers. Due to data availability at the time of writing, this paper does not analyze developments after this announcement of loosening, focusing instead on the initial calibration of the regulations.

Table 1. Overview of Macroprudential Regulations for Mortgage Lending

Loan-to-Value Limits	Primary Dwelling Homes	FTBs: Sliding LTV limit from 90 percent* SSBs: 80 percent
Loan-to-Income Limits	Primary Dwelling Homes	3.5 times LTI limit

*FTBs are allowed a 90 percent LTV up to a house value of €220,000. An 80 percent LTV applies above this value.

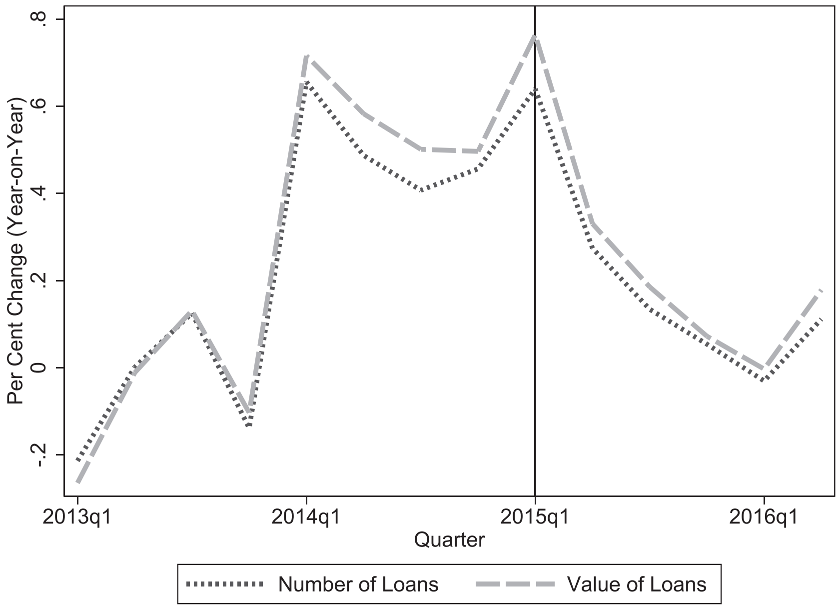
limit applies, with the first €220,000 of their purchase requiring a 10 percent deposit and any balance above €220,000 requiring a 20 percent deposit. For SSBs, a maximum LTV of 80 percent applies to all property purchases. Both borrower types are subject to an LTI limit of 3.5 times gross income. The regulations also allow for a certain value of new lending to exceed the limits, as it follows a proportional allowances model. Financial institutions are permitted to lend up to 15 percent of the value of new principal dwelling house (PDH) lending in excess of the LTV limit and up to 20 percent of the value of their new PDH lending in excess of the LTI limit.[3] Buy-to-let investors are subject to a 70 percent LTV cap.

Our main research hypothesis predicts a spillover of risk from the home to foreign markets as the Irish banks faced restrictions on risk-taking activity domestically. While it is not the objective of this paper to assess the direct impact of the introduction of mortgage measures in Ireland, it is important to demonstrate a material impact of the measures in Ireland on the activity of Irish banks. Our approach to this is twofold: first, we draw on aggregate data to demonstrate the overall impact on mortgage lending, and second, we draw on existing research to demonstrate the impact of the rules on the lending conditions (LTV and LTI ratios in the market).

The introduction of the regulations in Ireland occurred in February 2015 and were announced to the market in October 2014 when

[3]Self-builds are also subject to the regulations, as are refinances/switchers with an increase in capital. There are number of exemptions to the regulations for negative equity loans, mortgage resolution agreements, and refinances/switchers without an increase in capital.

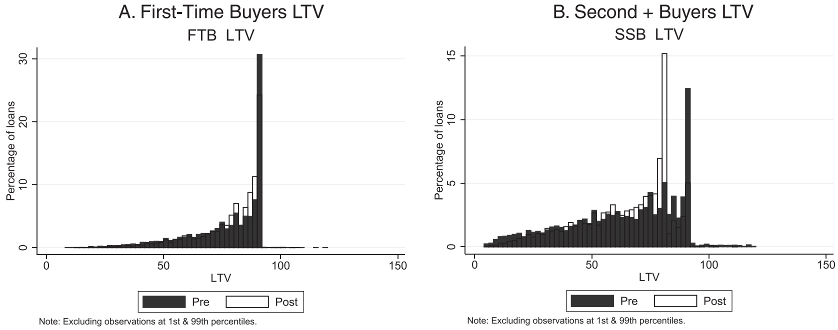
Figure 1. Growth in New Mortgages: ROI, 2013:Q1–2016:Q2



Source: Banking and Payments Federation Ireland.
Note: Vertical line indicates introduction of policy.

the central bank published research on their potential benefits (Hallissey, Kelly, and O'Malley 2014) as well as opening a public consultation on the proposed regulations. This move was not expected by the market, and their unanticipated announcement represented an exogenous change to the loan-to-value and loan-to-income limits applying to new mortgage loans to a market which had not had such a framework before. The overall result of the introduction of the measures was a reduction in the growth rate of mortgage loans. Figure 1 presents the growth in new mortgage drawdowns (number of loans and the value of lending) from the period 2013:Q1 to 2016:Q2 for the Irish market. The data are taken from the Banking and Payments Federation Ireland and mainly cover the five largest lenders in the Irish market: Allied Irish Banks (including EBS Building Society), Bank of Ireland, KBC, Permanent TSB, and Ulster

Figure 2. Distributional Comparisons of LTV by Buyer Type



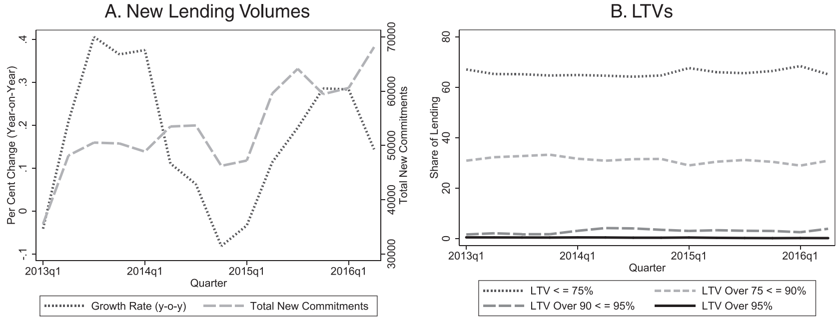
Source: Kinghan, McCarthy, and O'Toole (2017).
Notes: Pre-policy period: 2013:Q1–2014:Q3. Post-policy period: 2015:Q1–2016:Q2.

Bank. It can be seen that following the regulations, there was a marked slowdown in the growth of new lending.

To explore the impact of the regulations on the actual lending conditions in the market, we draw on two research papers that explore the changes in borrower leverage following the introduction of the measures. Kinghan, McCarthy, and O'Toole (2016) assess the LTV and LTI ratios in the market before and after the introduction of the regulations. They find that while on average there was actually a small rise in the indicators, for borrowers with a high demand for leverage,[4] there were falls in both the loan-to-income and loan-to-value ratios. For SSBs, the LTV for this group fell by up to 4 percentage points. Further research by Kinghan, McCarthy, and O'Toole (2017) highlights the fall for FTBs in LTV of up to 1.4 percentage points. Data in Kinghan, McCarthy, and O'Toole (2016) also highlight a statistically significant increase in incomes for FTBs in the market, which may be indicative of some change in borrower composition. Figure 2, reproduced from Kinghan, McCarthy, and O'Toole (2017), clearly demonstrates that the rules curbed the upper end of the LTV distribution, thus materially tightening the risk-taking of Irish banks.

[4]They define borrowers with a high demand for leverage as those with an LTV greater than or equal to 80 percent or an LTI greater than or equal to 3.

Figure 3. Overview of Activity in U.K. Mortgage Market



Source: Bank of England, MLAR Statistics.

2.2 The U.K. Mortgage Market and the Irish Connection

Mortgage debt in the United Kingdom accounts for approximately two-thirds of all outstanding loans (Bank of England 2016). While the Irish housing market suffered following the financial crisis, the effects in the United Kingdom were much less pronounced, with mortgage lending and house prices continuing to rise from 2013 onwards. A comparative description of house price developments for the United Kingdom and ROI is presented in appendix 1. Figure 3 presents an overview of activity in the U.K. mortgage market from 2013 to 2016 (the period under investigation in this paper). It can be seen that while the growth rates dropped in 2014, there has been a marked pickup more recently. In terms of risk-taking appetite in the United Kingdom, a majority of U.K. loans are issued at low LTVs (less than 75 percent) and, furthermore, there has been no discernible change in the level of LTVs on loans over the period.

Despite the relative stability in the mortgage market and the low level of mortgage delinquencies, in June 2014 the Financial Policy Committee of the Bank of England introduced a package of mortgage measures aimed at insuring against the risk of a marked loosening in underwriting standards in the owner-occupier mortgage market.[5]

[5]The Bank of England monitors delinquencies in the U.K. mortgage market in its annual Financial Stability Reports and the level is low. Kelly and O'Toole (2018) describe the level of delinquencies in the United Kingdom experienced by Irish lenders.

These measures included an affordability test and an LTI limit of 4.5 times income with a 15 percent allowance. The policies were not meant to be currently binding in the market, and a subsequent review of their functionality found they only had a very modest effect, with most lenders already applying more critical tests than the regulatory minimum (Bank of England 2016). This limit compares to a much stricter regime in Ireland, which may have left open the possibility for cross-border risk shifting such as that tested in this paper.

In terms of the presence of Irish lenders in the United Kingdom, there has been a long tradition of Irish involvement in direct lending in the United Kingdom which pre-dates Irish independence and continued thereafter (McGowan 1988). This was particularly strong in the case of Northern Ireland, with large subsidiaries of Irish banks continuing to operate north of the border. In the case of the banks present in our sample, Irish bank involvement in the U.K. market in the present day is not restricted to activity in Northern Ireland, but rather is distributed more generally across regions.

Presently, and of critical importance for this current paper, is that two main banking groups in Ireland (Bank of Ireland and AIB) have considerable ongoing activities in the U.K. market. Bank of Ireland has a particularly large presence in the United Kingdom. Its acquisition of Bristol and West Building Society in the 1990s gave it a foothold in the market and, currently, it lends directly as Bank of Ireland UK but also provides savings and loans products through its strategic alliance with the U.K. Post Office. Using data from its annual report, as of 2016:Q3, the total stock of U.K. mortgage loans outstanding for Bank of Ireland was valued at €23.8 billion, which is only marginally smaller than their Irish portfolio with €24.4 billion outstanding. In fact, at year-end 2016, the value of new mortgage lending was nearly two times larger in the United Kingdom, at €2.8 billion as compared to €1.4 billion in ROI.

For AIB, its U.K. presence is more regional, with a majority of activity being worked through First Trust, its wholly owned subsidiary in Northern Ireland. AIB also lends through AIB UK Group directly into the U.K. market. In total, AIB held €1.5 billion worth of U.K. mortgages outstanding.

3. Irish Banks' Lending at Home and Abroad

For our first empirical model, we use a simple difference-in-difference framework with loan-level data to compare the lending activity of Irish banks in the United Kingdom with their Irish lending before and after the introduction of macroprudential limits on LTV and LTI ratios in their Irish home market. This section documents the data and results for this comparative exercise.

3.1 Data and Identification Strategy

To investigate cross-border spillovers from macroprudential mortgage measures, we exploit loan-level data on the activity of Irish banks both domestically and abroad. The data come from regulatory loan-by-loan information that was submitted to the Central Bank of Ireland on a six-month basis since 2010. These data capture the universe of all outstanding loans for the Irish banks. These data were first provided as a requirement for the Prudential Capital Assessment Review (PCAR), which estimated the recapitalization requirements of the lenders following the financial crisis. The loan-level data (LLD) are currently collected from the five main lenders: Allied Irish Banks (including EBS), Bank of Ireland, KBC, Permanent TSB, and Ulster Bank.

For this particular paper, we draw on one specific drop of the LLD which includes all loans outstanding as of June 2016 for a subset of lenders with exposures in both the ROI and U.K. markets. In the ROI mortgage market, the loans covered account for around two-thirds of total mortgage market exposures. The data set is rich in capturing information that is recorded at origination for the purposes of credit risk assessment and lending allocation decisions. This includes loan characteristics including loan term, LTV, LTI, interest rate, interest rate type, drawn balance at origination, and loan type as well as borrower characteristics and property characteristics (price, region, dwelling type). The borrower characteristics are income (in euros or pounds sterling), age in years, employment status (employee, self-employed, other), and marital status (married, single, divorced, other). The property dwelling type takes the following categories: detached houses, semi-detached, terraced houses, and apartments.

Not all of the Irish banks have lending activity in the U.K. market. Requiring banks to have both U.K. and Irish lending leads us to construct a pooled cross-sectional data set of all loans issued by two Irish lenders who are active simultaneously in issuing new loans in both the ROI and U.K. mortgage markets during our target sample period. In order to create a balanced data set around the introduction of the macroprudential mortgage regulations in February 2015, we restrict the data to all loans issued between January 2013 and June 2016. Due to confidentiality constraints, we are not able to present summary statistics for the differential trends in lending volumes and credit conditions for both lenders in the U.K. and Irish markets, as would be normal in a setup like ours. However, we can consult the annual report of Bank of Ireland, the largest Irish-resident mortgage lender in the United Kingdom, which states the following:

> A key objective for 2015 was to continue to grow our mortgage business, building on the progress we made last year. In 2015, our new mortgage lending was GBP 3.3 billion compared with GBP 1.8 billion in 2014.

While clearly not causal, the above statement shows that, in aggregate terms, new lending expansion by the largest mortgage player was indeed occurring in the first year after the introduction of the domestic macroprudential mortgage restrictions in Ireland.

Table 2 presents summary statistics for the regression sample. Overall, there are 60,540 mortgages issued across the two jurisdictions by our sample banks between 2013:Q1 and 2016:Q2. The average LTI is 2.54 across all loans, while the average LTV ratio at origination is 72.04 percent. The average mortgage term is 290 months (24.17 years), while the average interest rate is 3.31 percent and the average borrower age is 37.57 years. The average unemployment rate across all regions and quarters is 7.37, while the average increase in the regional house price index from 2012:Q2 was 17.72 percent. In the second panel, we report the share of certain key categorical variables in the total sample. 41.9 percent of all mortgages were issued in Dublin, London, or their surrounding commuter regions (mideast in the ROI, southeast in the United Kingdom). 38.1 percent of mortgages were to first-time buyers, 27.2 percent to

Table 2. Summary Statistics for ROI-UK Regression Model

Variable	N	Mean	Std. Dev.
LTI	60,540	2.54	1.10
LTV (%)	60,540	72.04	18.92
Term	60,540	289.93	89.31
Interest Rate	60,540	3.31%	0.01
Age	60,540	37.57	8.32
Unemployment	60,540	7.37	2.47
House Price Index	60,540	117.72	17.10
Capital City or Commuter Belt	60,540	41.9%	
First-Time Buyer	60,540	38.1%	
Second and Subsequent Buyer	60,540	27.2%	
Refinance	60,540	24.7%	
Buy to Let	60,540	10.0%	
Salaried Employee	60,540	89.6%	

Notes: This table includes the summary statistics for the Irish lenders only used in the regression sample. LTI: loan-to-income ratio; LTV: loan-to-value ratio; Term: loan term in years. Capital city or commuter belt counties surround Dublin and London.

second and subsequent buyers, 24.7 percent for refinancing activity, and 10 percent for buy-to-let purchases. Finally, 89.6 percent of mortgages were to salaried employees, with the remainder going to self-employed borrowers.

Our main goal is to test how the lending standards of Irish banks changed in the United Kingdom following the regulations in Ireland, as measured by changes in LTV and LTI ratios. To this end, we estimate the following equation:

$$Y_{ijt} = \beta_1 UK + \beta_2 Post + \beta_3 (UK \times Post) + \beta_{4,5} X_{ijt} + \beta_{6,7} X_{rt} + \delta_{LoanGroup} + e_i, \tag{1}$$

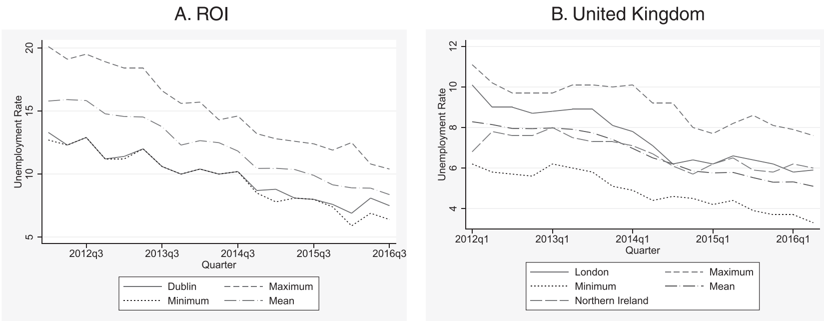
where Y_{ijt} is the LTV and LTI of mortgage i which was issued in country j in quarter t; X_{ijt} are a set of loan-level control variables which are not used to construct $\delta_{LoanGroup}$, such as the loan's interest rate and term at origination; and X_{rt} are regionally varying measures of economic performance which are likely to affect the demand for loans, such as the regional unemployment rate and the regional house price index, based to 100 at 2012:Q1 for each region across both jurisdictions.

$\delta_{LoanGroup}$ are a set of "loan-borrower group" fixed effects which saturate elements of loan-specific demand for debt which are likely to be common across borders, similar in spirit to those used by Auer and Ongena (2016). To construct these δ, we create buckets based on unique combinations of the following indicators for borrower types: three borrower age groups for borrowers aged eighteen to thirty-five, thirty-six to forty-five, and forty-six to sixty-nine; year-country house price quartiles; year-country income quartiles; an indicator for four property types (detached houses, semi-detached, terraced houses, apartments); and an indicator for capital cities and their surrounding commuter zones (taking a value of one for Dublin and the mideast regions in ROI, and London and the (outer) southeast regions in the United Kingdom). This leads to 2,382 groups that are included in the final model once we exclude all cells with one loan. The average number of loans in a cell is 351, with the median being 107 and the maximum being 3,720. Our empirical identification strategy relies on the assumption that, within each of these groups, any difference in LTV or LTI across borders over time which is not attributable to banks' supply decisions is controlled for in the X_{ijt} and the X_{rt}.

The crux of our empirical analysis is the coefficient on $(UK \times Post)$. The coefficient on UK will control for any differential in LTV or LTI between ROI and UK which is constant over the 2013–16 period; $Post$ will control for LTI and LTV differences that are common across jurisdictions in the period before and after February 9, 2015; and the term $(UK \times Post)$ is the difference-in-difference (DiD) coefficient that will capture the LTV or LTI differential across geographies which exists solely in the period after the introduction of the regulations, all the time controlling for the set of X_{ijt}, X_{rt}, and $\delta_{LoanGroup}$.

Unlike the majority of studies of the effect of policy interventions using DiD methods, careful attention must be paid to the fact that treated and control loans in our setting are located in different countries. As well as having different structural time-invariant features such as preferences for homeownership, enforceability of collateral,

Figure 4. Unemployment across the Two Jurisdictions: 2012 to 2016



and government housing market policy, loans in the treated and control groups are also likely to be facing different aggregate economic conditions during the period under study. To control for the possibility that such economic factors may introduce bias into our estimate of the treatment effect for U.K. loans issued after February 9, 2015, we include unemployment rates and house price indexes in our models, both of which vary at the region-quarter level. Figure 4 plots the maximum, minimum, and mean unemployment rate across Ireland's eight NUTS3 regions from 2012 to 2016, as well as selecting the Dublin rate for exposition. The right-hand panel does so similarly for the United Kingdom's fourteen NUTS3 regions, while selecting the London rate. The graphs show a clear improvement in economic performance in both countries over the period, with significant variation across regions within each country which can be controlled for in our empirical framework.

The standard difference-in-difference model relies on the assumption that while counterfactual levels of the treated and non-treated groups can be different, their counterfactual time variation must be similar. In the context of our current study, we need it to be the case that in the absence of macroprudential policy in the Republic of Ireland, the LTV and LTI on new loans, controlling for level differences in the outcome variable across the two jurisdictions and the set of X_{ijt}, X_{rt}, and $\delta_{LoanGroup}$, would have evolved in the same way.

While controlling for observables can do much of the work in ensuring that our β_3 is interpretable as the difference between ROI banks' U.K. risk-taking under the policy versus under the counterfactual no-policy scenario, issues can still remain in the case where repeated cross-sections with time-varying sample composition are being used, rather than panel data. Given that our data set reports the profile of newly originated mortgages, a layer of complexity is added, as we are not analyzing two groups of loans (treated and control, both of which were observed before and after the policy change) but four groups of different loans (pre-policy treated, pre-policy control, post-policy treated, post-policy control). This is precisely the type of pooled cross-sectional data where changing sample composition can hamper the interpretation of β_3 as a true DiD coefficient. To give a concrete example, if the share of first-time buyers in total originations has shifted dramatically between the pre- and post-policy period in ROI as a result of the policy, and we run an unweighted model to estimate β_3, it may be that the DiD coefficient is actually being driven by the changing sample composition, (with first-time buyers always being more likely to have higher-LTV loans), rather than by true, ceteris paribus changes in risk-taking for a given loan.

To address this issue, we use a propensity-score matching approach to match the borrowers on observables and then use these propensity scores to weight the DiD regression, proposed by Stuart et al. (2014). This involves estimation of a multinomial logit model to deal with the four-group nature of the selection issue. Using the estimated multinomial logit model, each observation in the data set gets a predicted probability of being in each of the groups, $Pr_g(X_i)$: (i) pre-policy treated, (ii) pre-policy control, (iii) post-policy treated, and (iv) post-policy control. The following inverse probability weights are then calculated:

$$w_i = \frac{Pr_1(X_i)}{Pr_g(X_i)}, \tag{2}$$

$$where\ 1 = \left(\sum_{g=1}^{4} Pr_g(X_i)\right) \forall i.$$

These weights can then be used in the DiD of equation (1) to account for changing sample composition. All regressions run for the remainder of this paper include weights derived from a multinomial logit model where a loan's position in the above four-category variable is explained by dummies for salaried employee, interest rate type (fixed versus variable), property type, within year-country house price quartiles, within year-country income quartiles, and buyer types (for the overall model).

3.2 Empirical Results

Table 3 provides estimates of equation (1) where Y_{ijt} is the LTV of mortgage i issued in country j (ROI or United Kingdom) in quarter t. The term $UK^{*}Post$ corresponds to β_3 from equation (1). The interpretation of the coefficient is that, controlling for a host of loan- and household-level controls which capture the composition of lending, as well as regional, time, and bank dummies, there is a 5.35 to 5.66 greater LTV on U.K. mortgages after the introduction of the regulations. This corresponds to close to one-third of a standard deviation of originating LTV in the estimation sample and is robust to the inclusion of a continuous and region-varying measure of unemployment (which is measured at quarterly frequency) and house price indexes in columns 2 and 3, both of which are expected to capture much of the local aggregate economic developments, which could explain diverging paths in LTV that are not due to the regulations.

We investigate where banks' risk-taking spillovers are confined to particular classes of borrower in columns 4–7. This breakdown suggests that the largest effect is found in the first-time buyer segment (a DiD coefficient of 6.88), where one would generally expect to find the most credit-constrained borrowers. A large and significant estimate of β_3 (an LTV differential of 5.78 points) is also found in the mover-purchaser (SSB) market. Among refinancing mortgages, the coefficient on β_3 is not statistically different from zero, whereas in the buy-to-let market U.K. loans in the post-period are estimated to have lower LTV, although this effect is insignificant at the 5 percent level.

Table 4 runs the same equation with the loan-to-income (LTI) ratio as the dependent variable. Overall, and in the SSB and refinancing segments, there is in fact a negative coefficient on β_3. This implies that, in the period after the introduction of macroprudential mortgage regulation in Ireland, LTIs on Irish banks' mortgages grew by a relatively larger amount in the ROI relative to the United Kingdom. These effects may possibly be explained by the fact that the

Table 3. Difference-in-Difference Model of Irish-Headquartered Banks' ROI and U.K. Lending: LTV

	(1) All	(2) All	(3) All	(4) FTB	(5) SSB	(6) Refi	(7) BTL
UK × *Post*	5.354***	5.661***	5.606***	6.882***	5.784***	1.564	−4.230*
	(0.786)	(0.845)	(0.839)	(1.185)	(1.223)	(3.152)	(2.165)
United Kingdom	6.952***	5.611***	6.217***	3.133*	4.938***	13.05***	12.21***
	(0.928)	(1.407)	(1.381)	(1.627)	(1.674)	(3.656)	(2.474)
Post	−1.699**	−2.275***	−1.793**	−1.338	−1.660	1.680	6.237***
	(0.721)	(0.872)	(0.853)	(1.259)	(1.284)	(3.157)	(2.164)
Loan Term	0.0873***	0.0869***	0.0870***	0.0695***	0.100***	0.0992***	0.0513***
	(0.00361)	(0.00356)	(0.00355)	(0.00484)	(0.00684)	(0.00738)	(0.00497)
Regional Unemployment		−0.303	−0.310*	−0.368	−0.0191	−0.356	0.237
		(0.185)	(0.187)	(0.270)	(0.206)	(0.400)	(0.263)
Regional HP			−0.0485*	−0.0255	−0.107***	−0.0378	0.0665
			(0.0250)	(0.0270)	(0.0274)	(0.0560)	(0.0426)
Observations	71,597	71,593	71,593	28,611	20,149	16,687	6,526
Loan Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mlogit PSM Weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors are in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. "Loan Group Fixed Effect" defines the unique combinations of borrower age groups, within-year house price quartile, within-year income quartile, and property type. In columns 1–3 it also distinguishes between first-time buyer, second and subsequent buyer, and refinance transactions. The model uses a linear specification estimated by OLS. Standard errors are robust to heteroskedasticity. Columns 4–7 estimate the regression specifically on the subsamples: FTB, SSB, Refi, and Buy-to-Let (BTL). The dependent variable is the loan-to-value ratio. Mlogit PSM weights are generated from multinomial logit specification as described in section 3.1.

Table 4. Difference-in-Difference Model of Irish-Headquartered Banks' ROI and U.K. Lending: LTI

	(1) All	(2) All	(3) All	(4) FTB	(5) SSB	(6) Refi	(7) BTL
UK × Post	−0.292***	−0.285***	−0.284***	0.303***	−0.217***	−0.643***	0.217***
	(0.0450)	(0.0506)	(0.0505)	(0.0449)	(0.0563)	(0.118)	(0.0815)
United Kingdom	0.587***	0.559***	0.542***	0.217***	0.598***	0.565***	−0.0861
	(0.0401)	(0.686)	(0.0680)	(0.0545)	(0.0776)	(0.139)	(0.113)
Post	0.281***	0.269***	0.255***	−0.123***	0.186***	0.410***	−0.0985
	(0.0370)	(0.0489)	(0.0490)	(0.0456)	(0.0564)	(0.112)	(0.0783)
Loan Term	0.00392***	0.0391***	0.00391***	0.00331***	0.00503***	0.00363***	0.00232***
	(0.000166)	(0.000166)	(0.000166)	(0.000195)	(0.000349)	(0.000317)	(0.000258)
Regional Unemployment		−0.00639	−0.00618	−0.0577***	−0.00232	0.00466	−0.0580***
		(0.00973)	(0.00976)	(0.00928)	(0.0113)	(0.0158)	(0.0139)
Regional HP			0.00131	0.00665***	0.00346**	−0.00374**	0.00929***
			(0.000999)	(0.00118)	(0.00168)	(0.00189)	(0.00281)
Observations	72,358	72,354	72,354	28,900	20,627	16,703	6,505
Loan Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mlogit PSM Weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors are in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. "Loan Group Fixed Effect" defines the unique combinations of borrower age groups, within-year house price quartile, within-year income quartile, and property type. In columns 1–3 it also distinguishes between first-time buyer, second and subsequent buyer, and refinance transactions. The model uses a linear specification estimated by OLS. Standard errors are robust to heteroskedasticity. Columns 4–7 estimate the regression specifically on the subsamples: FTB, SSB, Refi, and Buy-to-Let (BTL). The dependent variable is the loan-to-income ratio. Mlogit PSM weights are generated from multinomial logit specification as described in section 3.1.

LTI limits set by the regulation were comfortably above the typical LTI being issued in the Irish market in 2014, meaning that there was scope for Irish banks' ROI LTIs to grow substantially while remaining within the limits set by the regulations. This intuition is confirmed in the ROI market by Kinghan, McCarthy, and O'Toole (2016), who show that the average LTI went from 2.7 to 2.9 in the FTB segment of the Irish mortgage market after the introduction of the LTI limit in early 2015. The only borrower group in which the expected risk spillover is found is the first-time buyer group, in column 4. The estimated effect (an LTI increase of 0.3) in this case is one-third of a standard deviation in magnitude.

Our baseline empirical specifications do not include the mortgage interest rate as a control variable, on the basis that banks in general price their lending products as step functions of the loan-to-value ratio, with higher-leveraged loans paying an interest premium. Given this schedule, there is a concern that the coefficient on the interest rate in a model of LTV or LTI would be mechanically determined. However, in order to confirm that our DiD results are robust to the possibility that the interest rate is in fact an appropriate control variable to include in the empirical models, we repeat the specifications from tables 3 and 4 with the interest rate included in all cases. The results are reported in appendix 2 in tables 11 and 12, and in most cases the DiD coefficients are of similar sign, magnitude, and statistical significance. In particular, the positive spillover coefficient for LTV across the entire market as well as in the FTB and SSB segments is robust, as is the LTI coefficient in the FTB segment.

4. Comparing Irish and U.K. Banks' Activity in the U.K. Mortgage Market

As noted in the previous section, our difference-in-difference approach based on comparing the Irish banks' lending in both the United Kingdom and Ireland requires that, in the absence of macroprudential policy in the Republic of Ireland, the LTV and LTI on new loans, controlling for observables, would have evolved in the same manner (parallel trends). This is a strong assumption and could be violated if the lending trends in the United Kingdom were diverging from those observed in Ireland and this divergence happened to coincide with the introduction of limits on borrower measures in Ireland.

To establish that the Irish banks' risk-taking diverged from other banks in the U.K. market after the macroprudential limits in Ireland, the correct identification strategy would compare the risk-taking of "treated" Irish banks with that of a "control" group of other banks, both lending to similar borrowers in the U.K. market. This would ensure that the overall market trends and the Irish banks' idiosyncratic risk appetite are controlled for and any difference between the groups in the period after the Irish regulations can be attributed to a change in the Irish banks' behavior. In this section, we introduce data from other U.K. lenders and attempt to empirically test this proposition.

4.1 Data and Identification Strategy

To undertake an assessment within country for the United Kingdom, testing whether Irish banks' lending conditions changed following the introduction of macroprudential regulations in their home market, we append loan-by-loan data from a selection of U.K. lenders to our proprietary supervisory data on the Irish banks. The data are taken from the loan-by-loan information submitted to the European Data Warehouse as part of the asset-backed securities collateral reporting requirements of the Eurosystem. We capture data from securitized pools of U.K. mortgages that were submitted since 2013 and cover loans originated from January 2013 to December 2016. The data cover the following lending institutions: PSB, Skipton Building Society, TSB, and Virgin Money plc.

Using the data from these banks and the Irish data, we respecify our difference-in-difference equation as follows:

$$\begin{aligned} Y_{ibt} = {} & \beta_1 ROILenders + \beta_2 Post + \beta_3 (ROILenders \times Post) \\ & + \beta_4 \mathbf{X_{ijt}} + \gamma_{r,post} + \delta_{LoanGroup} + e_i, \end{aligned} \tag{3}$$

where Y_{ibt} are the LTV and LTI of mortgage i which was issued by bank b in period t. The β_1 parameter captures the difference in the LTI and LTV for Irish banks relative to U.K. banks as measured by the Irish bank identifier *ROI Lenders*. The β_2 parameter captures any difference between the pre- and post-regulation periods on average. Most importantly, the difference-in-difference parameter β_3

captures how much higher the Irish banks' LTV and LTI levels were in the post-regulation period.

As in equation (1), X_{ijt} is a vector of loan-level control variables including the interest rate type on the contract (fixed or adjustable), and the mortgage term at origination. To purge the model of considerable influences from the borrower-demand side, we also saturate the model with the "Loan Group" fixed effects, $\delta_{LoanGroup}$, which are documented previously.

In our cross-country difference-in-difference setting, we required variation at the country-time level for identification, which implied that we could not include interaction terms between regional and time dummies, as these would be collinear with the difference-in-difference parameter. This motivated the inclusion of the time-varying unemployment and house price indexes. However, in this specification, given that we have lending by both treatment and control banks in all regions and all time periods, we can include interactions of regional dummies and quarterly time dummies to purge any region-specific effects and confounding factors from the local economies at this level of variation. We therefore include interactions of regional dummies with the *Post* dummy to purge these confounding factors. As before, the model is estimated using the multinomial logit propensity-score weights.

One limitation of the control sample of other U.K. lenders is that the data are only taken from loans that were securitized and used as collateral for access to central bank funding. This has two limitations: (i) the loans do not cover the population of mortgages in the United Kingdom and (ii) the loans are not the full set of mortgages issued by these institutions during this time frame. The data purely cover the loans put into securitized pools by these banks. In this regard, any selection biases regarding the type of loans and the institutions that are covered by this process may affect the estimation. However, given that this is a difference-in-difference strategy, any time-invariant selection into securitized pools is controlled for. Furthermore, given the lack of any other loan-level data for the United Kingdom to do this comparison, our use of these data to address this research questions is still a major contribution relative to the existing literature.

The combined data set contains approximately 80,000 observations. Summary statistics for the overall sample of U.K. and Irish

Table 5. Summary Statistics for Within-U.K. Regression Model

Variable	N	Mean	Std. Dev.
LTV (%)	79,745	64.64	20.40
LTI	79,745	2.60	1.26
Interest Rate	76,793	2.86%	0.88
Term	79,745	21.5	7.6
UKH (East of England)	79,745	7.9%	
UKI (London)	79,745	15.9%	
UKJ (Southeast England)	79,745	11.6%	
UKK (Southwest England)	79,745	9.1%	
UKL (Wales)	79,745	4.3%	
UKM (Scotland)	79,745	6.4%	
Other Regions	79,745	44.8%	
First-Time Buyers	79,745	17.0%	
Second and Subsequent Buyers	79,745	50.4%	
Refinance	79,745	32.6%	

Notes: Statistics cover Irish lenders' U.K. activity and EDW U.K. bank lending data. LTI: loan-to-income ratio; LTV: loan-to-value ratio; Term: loan term in years.

banks are included in table 5. Due to confidentiality constraints relating to the underlying data, it is not possible to provide charts which split out parallel trends or show the observations before and after the policy intervention for the treated and control group, as would be standard in this context. The mean LTV was 64 percent and the mean LTI was 2.6. The highest share of lending took place in the London region (UKI), with 15.9 percent of loans issued. There is an "Other Regions" category which is constructed by the authors. A region is placed into this category where there is no variation in the difference-in-difference variable within this region, i.e., for a region to be included in the model with its own fixed effect, there must be lending by both an Irish and a U.K. bank, in both the pre- and post-regulation periods, in that region. After applying this criterion, 44.8 percent of observations are placed into the "Other Regions" category. Just over half of the borrowers in the sample were second and subsequent borrowers, while 32 percent were refinancing borrowers and approximately 17 percent were first-time buyers.

Table 6. Difference-in-Difference LTV Model: U.K. Lending of Irish Banks vs. U.K. Lending of Local Banks

	(1) Baseline	(2) FTB	(3) SSB	(4) Refinance
Post × ROI Lenders	1.1021**	3.1194***	1.4042*	−0.6643
	(0.383)	(0.430)	(0.638)	(0.774)
ROI Lenders	2.5019***	2.8897***	1.5531***	3.7255***
	(0.278)	(0.367)	(0.438)	(0.554)
Post	4.6204***	3.4955***	4.4738***	5.3810***
	(0.210)	(0.337)	(0.332)	(0.392)
Loan Term	1.1519***	0.5673***	1.2513***	1.4046***
	(0.024)	(0.027)	(0.041)	(0.043)
Observations	79,745	13,586	40,195	25,964
Of Which No. FE	714	192	192	192
Region * Post FE	Yes	Yes	Yes	Yes
Mlogit PSM Weights	Yes	Yes	Yes	Yes

Notes: Standard errors are in parentheses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. "Loan Group Fixed Effect" defines the unique combinations of borrower age groups, within-year house price quartile, within-year income quartile, and property type. In column 1 it also distinguishes between first-time buyer, second and subsequent buyer, and refinance transactions. The model uses a linear specification estimated by OLS. Standard errors are robust to heteroskedasticity. Mlogit PSM weights are generated from multinomial logit specification as described in section 3.1. Columns 2–4 estimate the regression specifically on the subsamples: FTB, SSB, Refi, and Buy-to-Let (BTL). The dependent variable is the loan-to-value ratio.

4.2 Empirical Findings

In this section, we present the main findings of the within-UK regressions for LTV and LTI. In both tables, column 1 contains the overall sample results, column 2 restricts the sample to first-time buyers (FTB), column 3 restricts the model to second and subsequent borrowers (SSBs), and column 4 estimates the model on the sample of refinancing borrowers. In all of these specifications, we include loan group fixed effects, region-time dummies, and loan term. All models are also weighted using the multinomial logit probability weights to deal with shifts in sample composition.

In column 1 of table 6, the β_3 estimate suggests that Irish banks' LTVs grew by 1.1 percentage points more than their U.K.

Table 7. Difference-in-Difference LTI Model: U.K. Lending of Irish Banks vs. U.K. Lending of Local Banks

	(1) Baseline	(2) FTB	(3) SSB	(4) Refinance
Post × *ROI Lenders*	0.0852***	0.4537***	−0.1410***	0.0070
	(0.023)	(0.033)	(0.034)	(0.044)
ROI Lenders	−0.0706***	−0.1147***	−0.0890***	0.0734*
	(0.016)	(0.028)	(0.022)	(0.033)
Post	−0.3610***	−0.3779***	−0.4628***	−0.1943***
	(0.013)	(0.024)	(0.019)	(0.022)
Loan Term	0.0591***	0.0365***	0.0561***	0.0769***
	(0.001)	(0.002)	(0.002)	(0.003)
Observations	79,745	13,586	40,195	25,964
Of Which No. FE	714	192	192	192
Region * Post FE	Yes	Yes	Yes	Yes
Mlogit PSM Weights	Yes	Yes	Yes	Yes

Notes: Standard errors are in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. "Loan Group Fixed Effect" defines the unique combinations of borrower age groups, within-year house price quartile, within-year income quartile, and property type. In column 1 it also distinguishes between first-time buyer, second and subsequent buyer, and refinance transactions. The model uses a linear specification estimated by OLS. Standard errors are robust to heteroskedasticity. Mlogit PSM weights are generated from multinomial logit specification as described in section 3.1. Columns 2–4 estimate the regression specifically on the subsamples: FTB, SSB, and Refi. The dependent variable is the loan-to-value ratio.

counterparts between the pre- and post-regulation periods. Columns 2 and 3 show that this effect was more than twice as large in the FTB as it was in the SSB segment (3.1 versus 1.4 LTV points). This provides further evidence that Irish banks may have targeted loosened credit conditions at FTBs in the post-regulation period in the United Kingdom, with their LTV expansion being shown here to be greater than competitor U.K. banks. This pattern complements the findings of the previous section, which showed that the Irish banks' U.K. LTVs expanded by more than their ROI LTVs.

Table 7 presents the results for spillovers along the LTI channel. In column 1 for the full sample, the parameter is positive and statistically significant at the 0.1 percent level. This suggests that, in the period following the introduction of macroprudential measures in the

Irish mortgage market, Irish banks loosened loan-to-income ratios for their U.K. loans relative to a sample of U.K. banks. The magnitude of the coefficient suggests a 0.085 percentage point increase in the loan-to-income ratio following the Irish regulations, which is small relative to the sample standard deviation in originating LTI of 1.26.

The DiD estimates in columns 2–4 suggest that the loosening occurred in the FTB segment, and there is no evidence that a similar loosening occurred for SSBs or refinance loans. The magnitude of the effect for FTB is large relative to the overall effect, at 0.45 (over one-third of a standard deviation in LTI).

As in the previous section, we repeat the specifications of tables 6 and 7 while including the interest rate as a control variable, to ensure that the results are robust to the importance of loan pricing in borrowers' LTV and LTI choices. Table 13 for LTV and table 14 for LTI (both reported in appendix 2) confirm that the direction, magnitude, and statistical significance of many of our spillover findings are maintained when including this control variable. In particular, the credit loosening for FTBs along both the LTV and LTI channel is robust, along with the overall LTV loosening.

5. A Focus on High-Risk Loans

We extend the analysis by moving away from estimations of linear models on changes in average LTV and LTI across groups, to focus on the probability of high-risk loans being issued rather than on average effects stemming from continuous variation in LTV and LTI. From a financial stability standpoint, the share of loans being issued in this right tail of the distribution may in fact be more relevant than changes in average LTI and LTV, given the likelihood that future defaults are concentrated in this region. In this section, we restrict our analysis to the U.K.-only models, comparing the propensity of ROI and local lenders to issue such high-risk loans before and after the policy introduction. Tables 8 and 9 present the results of these alternate specifications. For the dependent variable, we create a "high-LTV" dummy variable which takes a value of one above the LTV value of 85 percent (which covers 11.6 percent of loans issued in our U.K. data) and a "high-LTI" dummy which takes a value of

Table 8. Difference-in-Difference U.K. Lending Model: Probability of High-LTV Loan Being Issued

	(1) Baseline	(2) FTB	(3) SSB	(4) Refinance
Post × *ROI Lenders*	0.1020***	0.1911***	0.0606***	0.0541**
	(0.009)	(0.020)	(0.011)	(0.019)
ROI Lenders	0.0660***	0.1000***	0.0155*	0.1181***
	(0.007)	(0.017)	(0.008)	(0.014)
Post	0.0981***	0.1185***	0.0842***	0.1072***
	(0.005)	(0.014)	(0.005)	(0.010)
Loan Term	0.0144***	0.0191***	0.0067***	0.0211***
	(0.000)	(0.001)	(0.001)	(0.001)
Observations	79,745	13,586	40,195	25,964
Of Which No. FE	714	192	192	192
Region * Post FE	Yes	Yes	Yes	Yes
Mlogit PSM Weights	Yes	Yes	Yes	Yes

Notes: Standard errors are in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. "Loan Group Fixed Effect" defines the unique combinations of borrower age groups, within-year house price quartile, within-year income quartile, and property type. In column 1 it also distinguishes between first-time buyer, second and subsequent buyer, and refinance transactions. The model uses a linear specification estimated by OLS. Standard errors are robust to heteroskedasticity. Mlogit PSM weights are generated from multinomial logit specification as described in section 3.1. Columns 2–4 estimate the regression specifically on the subsamples: FTB, SSB, and Refi. The dependent variable is the dummy for loan-to-value > 85 percent.

one above a value of 4 (which covers 14.3 percent of our sample). A linear probability model is then run for both LTV and LTI in much the same format as the specifications of tables 6 and 7.

The results of table 8 suggest that, when measured in this discrete way, spillover effects are present in all columns 1–4. According to our DiD specification, Irish banks in the post-regulation period were 10 percentage points more likely than their U.K. counterparts to issue a high-LTV loan to a similar borrower, with this effect being of the magnitude of 19 percentage points in the FTB segment and 6 percentage points in the SSB segment (with all three effects being precisely estimated at 99 percent or better). For refinancing mortgages, the effect of 5.4 percent is statistically significant at the 5 percent level.

Table 9. Difference-in-Difference U.K. Lending Model: Probability of High-LTI Loan Being Issued

	(1) Baseline	(2) FTB	(3) SSB	(4) Refinance
Post × ROI Lenders	0.1028***	0.1817***	0.0452***	0.1038***
	(0.007)	(0.019)	(0.006)	(0.012)
ROI Lenders	−0.0379***	−0.0860***	−0.0215***	−0.0136
	(0.005)	(0.015)	(0.004)	(0.009)
Post	−0.1230***	−0.1977***	−0.0695***	−0.1460***
	(0.004)	(0.012)	(0.005)	(0.007)
Loan Term	0.0089***	0.0117***	0.0053***	0.0114***
	(0.000)	(0.001)	(0.000)	(0.001)
Observations	79,745	13,586	40,195	25,964
Of Which No. FE	714	192	192	192
Region * Post FE	Yes	Yes	Yes	Yes
Mlogit PSM Weights	Yes	Yes	Yes	Yes

Notes: Standard errors are in parentheses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. "Loan Group Fixed Effect" defines the unique combinations of borrower age groups, within-year house price quartile, within-year income quartile, and property type. In column 1 it also distinguishes between first-time buyer, second and subsequent buyer, and refinance transactions. The model uses a linear specification estimated by OLS. Standard errors are robust to heteroskedasticity. Mlogit PSM weights are generated from multinomial logit specification as described in section 3.1. Columns 2–4 estimate the regression specifically on the subsamples: FTB, SSB, and Refi. The dependent variable is the dummy for loan-to-income > 4.

Table 9 repeats the specification of table 8 for the loan-to-income channel. The probability of a high-LTI loan is shown to be higher for the Irish lenders in the post-regulation period in all models. The overall effect of a 10 percentage point higher probability is broken out into an 18 percentage point difference in the FTB segment and a 4.5 and 10.4 percentage point differential in the SSB and refinance segments, respectively.

A final specification looks at high-risk loans along both the LTV and LTI channels. In table 10, we define a high-risk loan to be a loan with both an LTI greater than 4 and an LTV above 85 percent. Running a similar model to the previous two tables with this dependent variable yields familiar results: for the first-time buyer segment, the probability of a high-risk loan is 27 percent higher for ROI banks in the post-period than for competitor domestic banks. Statistically

Table 10. Difference-in-Difference U.K. Lending Model: Probability of Combined High-LTI/High-LTV Loan Being Issued

	(1) Baseline	(2) FTB	(3) SSB	(4) Refinance
Post × ROI Lenders	0.1582***	0.2776***	0.1057***	0.1111***
	(0.010)	(0.020)	(0.013)	(0.020)
ROI Lenders	0.0297***	0.0256	−0.0056	0.0916***
	(0.007)	(0.017)	(0.008)	(0.015)
Post	−0.0157**	−0.0585***	0.0124	−0.0223
	(0.006)	(0.016)	(0.007)	(0.012)
Loan Term	0.0194***	0.0219***	0.0119***	0.0277***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	79,745	13,586	40,195	25,964
Of Which No. FE	714	192	192	192
Region * Post FE	Yes	Yes	Yes	Yes
Mlogit PSM Weights	Yes	Yes	Yes	Yes

Notes: High LTI is defined as above 4; high LTV is defined as above 85 percent. Standard errors are in parentheses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. "Loan Group Fixed Effect" defines the unique combinations of borrower age groups, within-year house price quartile, within-year income quartile, and property type. In column 1 it also distinguishes between first-time buyer, second and subsequent buyer, and refinance transactions. The model uses a linear specification estimated by OLS. Standard errors are robust to heteroskedasticity. Mlogit PSM weights are generated from multinomial logit specification as described in section 3.1. Columns 2–4 estimate the regression specifically on the subsamples: FTB, SSB, and Refi.

significant effects of 10.6 and 11 percent are found in the SSB and refinance segments of the market also.

As with previous regression tables, the results of tables 8 and 9 are shown to be robust to the inclusion of the interest rate control in tables 15 and 16 in appendix 2. In all cases, the DiD parameters do not change materially in either magnitude or significance depending on the inclusion or exclusion of r.

6. Conclusion

In this paper, using loan-level data on mortgages from Irish-headquartered banks active in both the Republic of Ireland and the United Kingdom, coupled with loan-level data on U.K.-only lenders, we test whether there was a change in risk-taking by Irish banks in the United Kingdom following the introduction of macroprudential regulations in Ireland. Our method deploys two separate difference-in-difference (DiD) estimations to test whether Irish banks loosened their LTV and LTI ratios abroad following a regulatory tightening of these lending standards at home.

We find evidence that, relative to the control group of local unaffected U.K. mortgage lenders, the probability of a high-risk loan (a loan with both a high LTI and high LTV) being issued by an Irish mortgage lender in the United Kingdom increased by 28 percent in the first-time homebuyer segment after the policy introduction. The effect is estimated to be smaller, yet sizable, at 11 percent in the mover-purchaser and refinance segments.

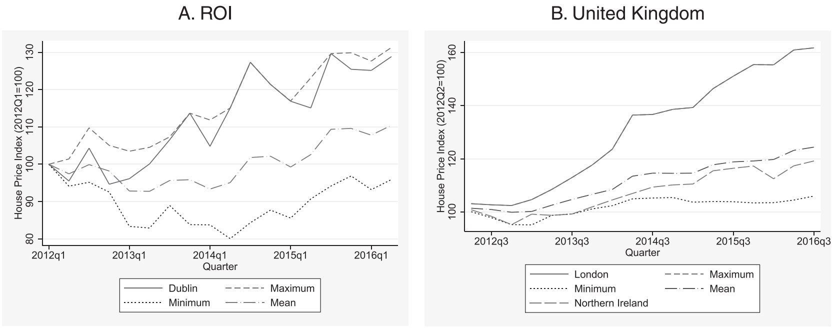
While our estimates show that risk-taking of Irish banks in the U.K. market was higher after the introduction of domestic restrictions on mortgage lending, particularly in the first-time buyer segment of the market, these findings do not imply that the policy instruments were ineffective. The stated aims of the macroprudential mortgage restrictions introduced in Ireland in February 2015 were not to improve banks' capital adequacy, but rather to ensure the resilience of Irish households to potential adverse developments in the housing market, as well as to dampen the procyclical relationship between credit and house prices in Ireland. The findings presented in this paper are consistent with the introduced policy meeting its stated aims. The implications of the results suggest that other capital-based regulatory tools that focus on the entire balance sheet of regulated entities may be warranted in conjunction with such product-specific mortgage measures in order to account for the potential for spillovers such as those identified here.

The potential need for international coordination in financial regulation has been recognized in academic and policy literature for quite some time. Referring to patterns of Japanese bank penetration into the U.S. banking market in the 1990s, Acharya (2003) suggests that uneven application of financial regulation (or "regulatory forbearance" in the terminology of the paper) may have provided Japanese banks with a competitive advantage in the ability to take risk. His paper points out that the presence of such variation in regulatory forbearance has the potential to undermine the international coordination efforts on other policy levers such as Basel capital adequacy ratios. Vinals and Nier (2014) provide an overview of the issues at play in the international spillover debate, with a focus on the potential for international spillovers given the highly interconnected nature of the global financial system. They also point out that, unlike in the case of the countercyclical capital buffer where Basel Committee members are expected to act in a reciprocal way once an individual country applies the buffer, many other macroprudential instruments such as LTV and LTI restrictions do not have provisions in place to allow for such international coordination. Our research presents further evidence of the need to understand banks' international activities and, where possible, to ensure reciprocity and coordination in the implementation of macroprudential policies. Without such coordination, risks may continue to build.

Appendix 1. Macroeconomic Context

Figure 5 indexes the house price index of each region to its level in 2012:Q1, and shows again that there has been significant variation in the evolution of house prices since the tail-end of the crisis period in both jurisdictions. In particular, the disproportionately strong house price growth of both capital cities is evident in both panels, while in Ireland's case the worst-performing region had not even reached its 2012:Q1 level by 2016:Q2. In the United Kingdom, the London house price index had grown by just over 60 percent during the period, while the slowest-growing region by comparison was less than 5 percent above its opening level. Such geographic disparities are an important part of the setting for our study, with the graphs indicating that it is crucial that time-varying regional developments are controlled for before our DiD model can be reliably interpreted.

Figure 5. Regional House Prices, Indexed to 2012:Q1: 2012 to 2016



Appendix 2. Robustness of Results to Including the Interest Rate as a Control Variable

Table 11. Difference-in-Difference Eatimator for Originating Loan to Value, All ROI and U.K. Mortgages Issued by Irish-Headquartered Banks from 2013:Q1 to 2016:Q2

	(1) All	(2) All	(3) FTB	(4) SSB	(5) Refi	(6) BTL	(7) Est7
United Kingdom	10.21***	8.106***	7.862***	4.433*	5.735**	16.47***	14.60***
	(1.867)	(2.416)	(2.395)	(2.549)	(2.486)	(5.792)	(2.666)
Post	−1.824*	−2.415**	−2.607**	−0.839	−3.706**	2.562	4.919**
	(0.960)	(1.082)	(1.094)	(1.823)	(1.770)	(3.891)	(2.483)
United Kingdom × *Post*	4.692***	5.028***	4.812***	4.186**	5.396***	−0.671	−1.258
	(1.007)	(1.054)	(1.073)	(1.755)	(1.773)	(3.695)	(2.539)
Loan Term	0.0784***	0.0777***	0.0768***	0.0631***	0.0901***	0.0864***	0.0532***
	(0.00481)	(0.00464)	(0.00448)	(0.00642)	(0.00984)	(0.00858)	(0.00481)
Interest Rate	430.4***	430.5***	437.0***	486.9***	522.8***	477.7***	481.5***
	(23.59)	(23.31)	(23.03)	(27.08)	(46.15)	(45.17)	(58.66)
Regional Unemployment		−0.551	−0.480	−0.425	−0.214	−0.828	0.0355
		(0.368)	(0.344)	(0.388)	(0.365)	(0.794)	(0.244)
Regional HP			0.0641	0.00423	−0.0210	0.128	0.0131
			(0.0532)	(0.0189)	(0.0334)	(0.112)	(0.0250)
Observations	60,915	60,911	60,911	23,175	16,598	15,051	6,212
Loan Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. "Loan Group Fixed Effect" defines the unique combinations of borrower age groups, year-country house price quartile, year-country income quartile, property type, and an indicator for capital cities and their surrounding commuter zones. In columns 1–3 it also distinguishes between first-time buyer, second and subsequent buyer, refinance, and buy-to-let purchases.

Table 12. Difference-in-Difference Estimator for Originating Loan to Income, All ROI and U.K. Mortgages Issued by Irish-Headquartered Banks from 2013:Q1 to 2016:Q2

	(1) All	(2) All	(3) All	(4) FTB	(5) SSB	(6) Refi	(7) BTL
United Kingdom	0.472***	0.336***	0.309***	−0.00532	0.376***	0.0695	−0.120
	(0.0810)	(0.113)	(0.111)	(0.0899)	(0.116)	(0.266)	(0.136)
Post	0.226***	0.188***	0.167***	−0.279***	0.0221	0.290	0.0538
	(0.0572)	(0.0642)	(0.0646)	(0.0770)	(0.0748)	(0.205)	(0.0955)
United Kingdom × *Post*	−0.182***	−0.161**	−0.184***	0.391***	−0.0816	−0.482**	0.0971
	(0.0617)	(0.0656)	(0.0647)	(0.0731)	(0.0751)	(0.208)	(0.0993)
Loan Term	0.00376***	0.00372***	0.00362***	0.00292***	0.00516***	0.00346***	0.00225***
	(0.000277)	(0.000274)	(0.000268)	(0.000409)	(0.000636)	(0.000399)	(0.000264)
Interest Rate	−18.76***	−18.74***	−18.03***	0.567	−13.07***	−47.11***	−17.46***
	(1.908)	(1.901)	(1.817)	(1.476)	(3.652)	(3.155)	(4.810)
Regional Unemployment		−0.0357**	−0.0279**	−0.0846***	−0.0111	−0.0365	−0.0503***
		(0.0144)	(0.0138)	(0.0137)	(0.0166)	(0.0255)	(0.0140)
Regional HP			0.00698***	0.0132***	0.00925***	0.00224	0.00617***
			(0.00149)	(0.00146)	(0.00167)	(0.00342)	(0.00184)
Observations	61,434	61,430	61,430	23,352	16,958	15,059	6,185
Loan Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors are in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. “Loan Group Fixed Effect” defines the unique combinations of borrower age groups, year-country house price quartile, year-country income quartile, property type, and an indicator for capital cities and their surrounding commuter zones. In columns 1–3 it also distinguishes between first-time buyer, second and subsequent buyer, refinance, and buy-to-let purchases.

Table 13. Within-U.K. Difference-in-Difference Model: Loan-to-Value Ratios

	(1) Baseline	(2) FTB	(3) SSB	(4) Refinance
ROI Lenders	2.9460***	3.1703***	2.6185***	3.4490***
	(0.283)	(0.334)	(0.455)	(0.547)
Post	4.5545***	3.3529***	4.4687***	5.2860***
	(0.207)	(0.315)	(0.329)	(0.388)
ROI Lenders × Post	0.1619	1.9254***	0.3038	−1.1627
	(0.366)	(0.390)	(0.590)	(0.749)
Current Interest Rate	2.8149***	4.6687***	2.5625***	2.1300***
	(0.152)	(0.113)	(0.245)	(0.286)
Loan Term	1.1154***	0.4258***	1.2340***	1.3860***
	(0.025)	(0.026)	(0.041)	(0.044)
Observations	79,745	13,586	40,195	25,964
Loan Group FE	Yes	Yes	Yes	Yes
Of Which No. FE	714	192	192	192
Region * Post FE	Yes	Yes	Yes	Yes
Mlogit PSM Weights	Yes	Yes	Yes	Yes

Notes: Standard errors are in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 14. Within-U.K. Difference-in-Difference Model: Loan-to-Income Ratios

	(1) Baseline	(2) FTB	(3) SSB	(4) Refinance
ROI Lenders	-0.0927^{***}	-0.1139^{***}	-0.2090^{***}	0.0879^{**}
	(0.016)	(0.028)	(0.022)	(0.032)
Post	-0.3578^{***}	-0.3782^{***}	-0.4622^{***}	-0.1893^{***}
	(0.013)	(0.024)	(0.019)	(0.022)
ROI Lenders × Post	0.1320^{***}	0.4506^{***}	-0.0171	0.0330
	(0.021)	(0.033)	(0.031)	(0.042)
Current Interest Rate	-0.1401^{***}	0.0122	-0.2886^{***}	-0.1114^{***}
	(0.009)	(0.010)	(0.013)	(0.015)
Loan Term	0.0609^{***}	0.0361^{***}	0.0581^{***}	0.0779^{***}
	(0.001)	(0.002)	(0.002)	(0.003)
Observations	79,745	13,586	40,195	25,964
Loan Group FE	Yes	Yes	Yes	Yes
Of Which No. FE	714	192	192	192
Region * Post FE	Yes	Yes	Yes	Yes
Mlogit PSM Weights	Yes	Yes	Yes	Yes

Notes: Standard errors are in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 15. Difference-in-Difference U.K. Lending Model: Probability of High-LTV Loan Being Issued

	(1) Baseline	(2) FTB	(3) SSB	(4) Refinance
ROI Lenders	0.0849***	0.1141***	0.0438***	0.1063***
	(0.007)	(0.015)	(0.008)	(0.014)
Post	0.0953***	0.0113***	0.0840***	0.1032***
	(0.005)	(0.013)	(0.005)	(0.010)
ROI Lenders × Post	0.0620***	0.1310***	0.0314**	0.0327
	(0.009)	(0.018)	(0.010)	(0.019)
Current Interest Rate	0.1198***	0.2348***	0.0681***	0.0914***
	(0.004)	(0.005)	(0.005)	(0.007)
Loan Term	0.0128***	0.0120***	0.0063***	0.0203***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	79,745	13,586	40,195	25,964
Loan Group FE	Yes	Yes	Yes	Yes
Of Which No. FE	714	192	192	192
Region * Post FE	Yes	Yes	Yes	Yes
Mlogit PSM Weights	Yes	Yes	Yes	Yes

Notes: High LTV is defined as above 85 percent. Standard errors are in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. "Loan Group Fixed Effect" defines the unique combinations of borrower age groups, within-year house price quartile, within-year income quartile, and property type. In column 1 it also distinguishes between first-time buyer, second and subsequent buyer, and refinance transactions.

Table 16. Difference-in-Difference U.K. Lending Model: Probability of High-LTI Loan Being Issued

	(1) Baseline	(2) FTB	(3) SSB	(4) Refinance
ROI Lenders	-0.0387^{***}	-0.0867^{***}	-0.0245^{***}	-0.0123
	(0.005)	(0.015)	(0.004)	(0.008)
Post	-0.1229^{***}	-0.1974^{***}	-0.0695^{***}	-0.1455^{***}
	(0.004)	(0.012)	(0.005)	(0.007)
ROI Lenders × Post	0.1046^{***}	0.1845^{***}	0.0482^{***}	0.1062^{***}
	(0.007)	(0.019)	(0.006)	(0.012)
Current Interest Rate	-0.0055^{**}	-0.0107	-0.0072^{***}	-0.0101^{**}
	(0.002)	(0.006)	(0.002)	(0.003)
Loan Term	0.0090^{***}	0.0120^{***}	0.0054^{***}	0.0115^{***}
	(0.000)	(0.001)	(0.000)	(0.001)
Observations	79,745	13,586	40,195	25,964
Loan Group FE	Yes	Yes	Yes	Yes
Of Which No. FE	714	192	192	192
Region * Post FE	Yes	Yes	Yes	Yes
Mlogit PSM Weights	Yes	Yes	Yes	Yes

Notes: High LTI is defined as above 4. Standard errors are in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. "Loan Group Fixed Effect" defines the unique combinations of borrower age groups, within-year house price quartile, within-year income quartile, and property type. In column 1 it also distinguishes between first-time buyer, second and subsequent buyer, and refinance transactions.

Table 17. Difference-in-Difference U.K. Lending Model: Probability of Combined High-LTV/High-LTI Loan Being Issued

	(1) Baseline	(2) FTB	(3) SSB	(4) Refinance
ROI Lenders	0.0449***	0.0364*	0.0196*	0.0825***
	(0.007)	(0.016)	(0.009)	(0.015)
Post	−0.0180**	−0.0639***	0.0123	−0.0254*
	(0.006)	(0.015)	(0.007)	(0.011)
ROI Lenders × Post	0.1262***	0.2319***	0.0796***	0.0947***
	(0.010)	(0.019)	(0.012)	(0.020)
Current Interest Rate	0.0959***	0.1788***	0.0607***	0.0699***
	(0.004)	(0.005)	(0.006)	(0.007)
Loan Term	0.0181***	0.0165***	0.0114***	0.0271***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	79,745	13,586	40,195	25,964
Loan Group FE	Yes	Yes	Yes	Yes
Of Which No. FE	714	192	192	192
Region * Post FE	Yes	Yes	Yes	Yes
Mlogit PSM Weights	Yes	Yes	Yes	Yes

Notes: High LTI is defined as above 4. Standard errors are in parentheses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. “Loan Group Fixed Effect” defines the unique combinations of borrower age groups, within-year house price quartile, within-year income quartile, and property type. In column 1 it also distinguishes between first-time buyer, second and subsequent buyer, and refinance transactions.

References

Acharya, V. V. 2003. "Is the International Convergence of Capital Adequacy Regulation Desirable?" *Journal of Finance* 58 (6): 2745–82.

Adelino, M., A. Schoar, and F. Severino. 2016. "Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class." *Review of Financial Studies* 29 (7): 1635–70.

Aiyar, S., C. W. Calomiris, J. Hooley, Y. Korniyenko, and T. Wieladek. 2014. "The International Transmission of Bank Capital Requirements: Evidence from the UK." *Journal of Financial Economics* 113 (3): 368–82.

Aiyar, S., C. W. Calomiris, and T. Wieladek. 2014. "Does Macro-Prudential Regulation Leak? Evidence from a UK Policy Experiment." *Journal of Money, Credit and Banking* 46 (s1): 181–214.

Auer, R., and S. Ongena. 2016. "The Countercyclical Capital Buffer and the Composition of Bank Lending." BIS Working Paper No. 593 (December).

Bank of England. 2016. "The FPC's Review of Its 2014 Mortgage Market Recommendations." Financial Stability Report (November).

Basten, C., and C. Koch. 2015. "Higher Bank Capital Requirements and Mortgage Pricing: Evidence from the Countercyclical Capital Buffer (CCB)." Working Paper No. 26 (June), Institute of Economic Research, Hitotsubashi University.

Buch, C. M., and L. S. Goldberg. 2016. "Cross-border Prudential Policy Spillovers: How Much? How Important? Evidence from the International Banking Research Network." Staff Report No. 801, Federal Reserve Bank of New York (November).

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

Cizel, J., J. Frost, A. Houben, and P. Wierts. 2016. "Effective Macroprudential Policy: Cross-sector Substitution from Price and Quantity Measures." DNB Working Paper No. 498.

De Jonghe, O., H. Dewachter, K. Mulier, S. Ongena, and G. Schepens. 2016. "Some Borrowers are More Equal than Others: Bank Funding Shocks and Credit Reallocation." Working Paper.

Demyanyk, Y., and E. Loutskina. 2016. "Mortgage Companies and Regulatory Arbitrage." *Journal of Financial Economics* 122 (2): 328–51.

Duca, J. V., J. Muellbauer, and A. Murphy. 2010. "Housing Markets and the Financial Crisis of 2007-2009: Lessons for the Future." *Journal of Financial Stability* 6 (4): 203–17.

———. 2011. "House Prices and Credit Constraints: Making Sense of the US Experience." *Economic Journal* 121 (552): 533–51.

Engelhardt, G. 1994. "House Prices and the Decision to Save for Down Payments." *Journal of Urban Economics* 36 (2): 209–37.

European Systemic Risk Board. 2016. "A Review of Macroprudential Policies in the EU in 2015." Technical Report.

Hallissey, N., R. Kelly, and T. O'Malley. 2014. "Macro-prudential Tools and Credit Risk of Property Lending at Irish Banks." Economic Letter No. 10/EL/14, Central Bank of Ireland.

Houston, J. F., C. Lin, and Y. Ma. 2012. "Regulatory Arbitrage and International Bank Flows." *Journal of Finance* 67 (5): 1845–95.

International Monetary Fund. 2013. "Key Aspects of Macroprudential Policy." Background Paper.

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

Kelly, R., and C. O'Toole. 2018. "Mortgage Default, Lending Conditions and Macroprudential Policy: Loan-level Evidence from UK Buy-to-Lets." *Journal of Financial Stability* 36 (June): 322–35.

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

Kinghan, C., Y. McCarthy, and C. O'Toole. 2016. "The Effects of Macroprudential Policy on Borrower Leverage." Economic Letter No. 08/EL/16, Central Bank of Ireland.

———. 2017. "Leverage, Macroprudential Policy and Borrower Heterogeneity." Mimeo, Technical Research, Central Bank of Ireland.

McGowan, P. 1988. "Money and Banking in Ireland: Origins, Development and Future." *Journal of the Statistical and Social Inquiry Society of Ireland* XXVI (I): 45–137.

Ongena, S., A. Popov, and G. F. Udell. 2013. "'When the Cat's Away the Mice Will Play': Does Regulation at Home Affect Bank

Risk-taking Abroad?” *Journal of Financial Economics* 108 (3): 727–50.

Reinhardt, D., and R. Sowerbutts. 2015. “Regulatory Arbitrage in Action: Evidence from Banking Flows and Macroprudential Policy.” Working Paper No. 546, Bank of England.

Stuart, E. A., H. A. Huskamp, K. Duckworth, J. Simmons, Z. Song, M. E. Chernew, and C. L. Barry. 2014. “Using Propensity Scores in Difference-in-Differences Models to Estimate the Effects of a Policy Change.” *Health Services and Outcomes Research Methodology* 14 (4): 166–82.

Tripathy, J. 2017. “Cross-border Effects of Regulatory Spillovers: Evidence from Mexico.” Working Paper No. 684, Bank of England.

Vandenbussche, J., U. Vogel, and E. Detragiache. 2015. “Macroprudential Policies and Housing Prices: A New Database and Empirical Evidence for Central, Eastern, and Southeastern Europe.” *Journal of Money, Credit and Banking* 47 (S1): 343–77.

Vinals, J., and E. Nier. 2014. “Collective Action Problems in Macroprudential Policy and the Need for International Coordination.” *Financial Stability Review* (Banque de France) April (18): 39–46.