IJCB INTERNATIONAL JOURNAL OF CENTRAL BANKING Volume 17, Number 2 June 2021

Volume 17, Number 2 June 2021

IJCB

INTERNATIONAL JOURNAL OF CENTRAL BANKING

IJCB International Journal of Central Banking

Volume 17, Number 2 June 2021

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

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

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

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

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

ISSN: 1815-4654

Broadening Narrow Money: Monetary Policy with a Central Bank Digital Currency*

Jack Meaning,[a] Ben Dyson,[b] James Barker,[a]
and Emily Clayton[a]
[a]Bank of England
[b]Bank for International Settlements Innovation Hub

This paper discusses central bank digital currency (CBDC) and its potential impact on the monetary transmission mechanism. We first offer a general definition of CBDC which should make the concept accessible to a wide range of economists and policy practitioners. We then investigate how CBDC could affect the various stages of transmission, from markets for central bank money to the real economy. We conclude that monetary policy would be able to operate much as it does now, by varying the price or quantity of central bank money. Transmission may even be strengthened for a given change in policy instruments.

JEL Codes: E42, E52, E58.

1. Introduction

The feasibility and desirability of central banks issuing their own fiat versions of digital currencies has been the focus of a growing debate in recent years. Numerous central banks around the world are researching the topic, including the Bank of Canada (2017), the European Central Bank (Lagarde 2020), the People's Bank of China (2020), the Sveriges Riksbank (2021), and the Bank of England (2020). The policy community is being aided by the wider academic community—for instance, Bjerg (2018), Bordo and Levin (2019),

*We would like to thank John Barrdear, Andrew Hauser, Charles Kahn, Michael Kumhof, and Andrew Levin for a number of informative discussions. We also thank two anonymous referees for their helpful comments. The views here represent those of the authors only and should not be interpreted as an official position of the Bank of England in any way. Corresponding author e-mail: jack.meaning@bankofengland.co.uk.

Eichengreen (2019), and projects such as the Massachusetts Institute of Technology's Digital Currency Initiative. Despite this, the literature remains in relative infancy, with consensus around some fundamental issues only slowly beginning to form. What is more, the very concept of a central bank digital currency can appear inaccessible to economists who are not familiar with blockchain, distributed ledger technology, or similar terminology.

A first contribution of this paper is therefore to offer a general definition of central bank digital currency (CBDC) around which consensus can build and which, importantly, should feel familiar to anyone with a basic understanding of the monetary process. We define central bank digital currency as *any electronic, fiat liability of a central bank that can be used to settle payments, or as a store of value.*

By this definition, a CBDC must be both electronic and a central bank liability. But beyond this, there are important policy choices to be made around key design parameters which are under the control of the central bank. These include the range of counterparties that have access to CBDC; whether CBDC will be freely convertible for other kinds of central bank money and/or commercial bank deposits; and whether CBDC is interest bearing. We explore each of these parameters in more detail in this paper. The optimal setting of each will depend on what the policymaker is seeking to achieve with CBDC, and so is intrinsically linked to their policy objectives.

Under the general definition given above, CBDC already exists in the form of central bank reserves (which are electronic liabilities of the central bank). CBDC could therefore be viewed simply as electronic narrow money. But access to central bank reserves is usually granted to a limited range of counterparties, primarily banks. Introducing a universally accessible CBDC is conceptually equivalent to broadening access to central bank reserves to the widest possible range of domestic counterparties, including nonfinancial businesses and households.

This broadening of narrow money is not without potential complications. Banks are currently given privileged access to central bank money because they play a key role in both the payments system and the transmission of monetary policy. Importantly, central banks use access to central bank reserves, the interest paid on those reserves, and the creation of new reserves as fundamental tools of monetary policy. Consequently, it is important to analyze the effects of widening access to electronic central bank money on the implementation and transmission of monetary policy.

This paper explores the impact of introducing CBDC on monetary policy. After discussing a range of potential design parameters, we focus on a specific form of central bank digital currency which is universally accessible, interest bearing, and freely convertible for other forms of central bank money and commercial bank deposits. We explore how CBDC might affect the implementation of monetary policy decisions and the channels, speed, and strength with which those decisions are transmitted to the real economy, focusing on two key tools in the current monetary policy toolkit: (i) changes in policy rates and (ii) asset purchases (in the style of quantitative easing).

While it is difficult to make definitive, quantitative statements at this stage about the impact of CBDC, our analysis leads us to the broad conclusion that a universally accessible, interest-bearing, freely convertible CBDC could be used for monetary policy purposes in much the same way that central bank reserves are now. On the margin, there may even be reason to believe that the monetary transmission mechanism would be stronger for a given change in policy instruments.

The rest of the paper is laid out as follows. In section 2 we elaborate on our general definition of CBDC. In section 3 we discuss the various design choices that a central bank can make, and specify the benchmark model that will be the basis for our analysis in later sections. Sections 4 and 5 discuss the likely impact of CBDC on the monetary transmission mechanism, from the control of CBDC rates by the central bank to the channels through which this affects the real economy. Section 6 considers central bank asset purchases with CBDC. Section 7 briefly discusses some of the key financial stability risks introduced by CBDC, and section 8 concludes.

2. Defining Central Bank Digital Currency

The term *central bank digital currency* is currently used to refer to a wide range of potential designs and policy choices, with no single commonly agreed definition. This is, in part, due to the fact that the concept sits at the nexus of a number of different areas of research and brings together topics as diverse as computer science, cryptography, payments systems, banking, monetary policy, and financial stability. As a result, the debate around CBDC may at times appear opaque to economists who are not familiar with cryptocurrencies, blockchain, and distributed ledger technology (DLT).

As a response, we offer a general definition that is both an accurate representation of CBDC and stated in terms that will be familiar to any economist with a basic understanding of the monetary process. It is our hope that this will remove any perceived mystique around CBDC and promote wider discussion of CBDC among economists.

To that end, we define a central bank digital currency simply as *an electronic, fiat liability of a central bank that can be used to settle payments or as a store of value.* It is in essence electronic central bank—or "narrow"—money.

This definition sits within a number of wider definitions of money and digital assets. Bordo and Levin (2017) define a digital currency as "an asset stored in electronic form that can serve essentially the same function as physical currency, namely facilitating payments transactions." Our definition simply adds to this the condition that the asset in question is a fiat liability of the central bank. It can also be viewed within the framework of the "money flower" of Bech and Garratt (2017) that looks to define and classify a range of money-like assets according to four characteristics: electronic; central bank issued; peer-to-peer exchange; and universally accessible. Within the money flower representation, our definition of CBDC is depicted by the four core segments incorporating both retail and wholesale central bank cryptocurrencies, settlement or reserve accounts with the central bank, and deposited currency accounts.[1]

Beyond two inherent characteristics (i.e., that CBDC must be both electronic and a central bank liability), we leave our definition of CBDC deliberately general. This is because central banks must make a number of design choices which could give rise to different forms of CBDC (Bank for International Settlements 2018). While there remains considerable uncertainty around the optimal

[1]The analysis within Bech and Garratt (2017) goes on to focus on the subset of money that is based on cryptographic technology, and so discusses central bank *crypto*currencies—CBCC rather than CBDC.

setting of these design parameters, a general definition allows for a structured discussion without limiting the possibilities open to policymakers. Key policy choices that we will consider below include the range of counterparties that have access to CBDC; whether CBDC is freely convertible for other kinds of central bank money and/or commercial bank deposits; and whether CBDC is interest bearing.

This general definition also provides a framework by which any future research can define the subset of CBDC to which it refers and assess how its conclusions may vary were those parameters to be set differently. We therefore see this as an important contribution of our paper.

By this definition, one form of CBDC already exists and sits at the heart of the monetary system in most economies: central bank reserves are both electronic account balances and a fiat liability of the central bank. A number of studies, including He et al. (2017), implicitly assume the same in their analysis.

3. Design Characteristics of CBDC

The optimal design of a specific model of CBDC will depend on a central bank's motivation for introducing it. In central bank discussions to date, these motivations fall into four categories. First, CBDC could support monetary policy by strengthening existing tools or introducing new ones. Second, CBDC may have financial stability benefits, by serving as a risk-free asset for those who are not covered by deposit insurance, or by encouraging greater settlement in central bank money and eliminating settlement and credit risk from the payment system. Third, in economies where the use of cash is rapidly declining, such as Sweden, CBDC can be seen as a way of preserving public access to central bank money and safeguarding a critical national payments system. Finally, in countries where a higher proportion of the population is unbanked, CBDC may increase financial inclusion by providing universal access to electronic payments. The following parameters must be set according to these motivations and the central bank's wider objectives.

3.1 Access

One of the key parameters that policymakers would have to decide on is the range of counterparties that have access to CBDC. A CBDC may be universally accessible—in other words, it can be held by anyone for any purpose—or access may be restricted to a limited subset of economic agents, or for limited purposes.

Some authors—for instance, Fung and Halaburda (2016) and Bjerg (2017)—identify universal access as a fundamental characteristic of any CBDC. This view derives from a definition of *currency* as something that must be able to be held by anyone. However, it is possible that central banks may issue a CBDC that is available only to a subsector of the economy, such as a "retail CBDC" for households and nonfinancial businesses only, or a "wholesale CBDC" which can be used as a settlement asset in financial markets by firms that do not currently have access to central bank reserves (Bech and Garratt 2017). The question is then a semantic one about whether a CBDC that is *not* universally accessible could still be considered a central bank digital *currency*. In fact, the European Central Bank initially chose to use the broader term "digital base *money*" (Lagarde 2020), which we find to be more accurate, but for consistency we will stick with the more widely used term "CBDC."

The optimal degree of access will depend on the policy objective of the central bank. For example, central banks that introduce CBDC for financial inclusion objectives would be best to extend access as widely as possible, especially among households. However, if the purpose of CBDC is to ensure a safe and efficient payments system, consistent with the objectives of many central banks, then it could be argued that a CBDC with narrow access, such as the reserves systems employed currently in many economies around the world, is sufficient, especially in economies with efficient and secure banking and financial infrastructure. Since the financial crisis of 2007–09 a number of central banks, including the Bank of England (Bank of England 2015), have undertaken projects to widen access to central bank money to broker-dealers, central counterparties, and others, in order to improve financial stability. This suggests that if financial stability were the motivation for introducing CBDC, then the optimal level of access would include a wider range of financial institutions than just the banking system, but would probably fall short of universal access.

3.2 Interest Bearing

A second important policy choice is whether CBDC is interest bearing. An interest-bearing CBDC might pay positive, zero, or even negative rates at various points in the economic cycle. The arguments in favor of paying interest on short-term central bank liabilities have their origins in Friedman (1960), who argued that they should pay a rate of interest equivalent to the risk-free rate to ensure that the opportunity cost of holding money is equal to the marginal cost of its production. A CBDC with a variable interest rate could be used as the main instrument of monetary policy, set in order to guide the macroeconomy and to stabilize inflation and output.

A central bank may want to pay different rates to different types of holders of CBDC. For instance, it may wish to pay different rates to banks and nonbanks (including households), because banks play a special role in the transmission of monetary policy and the economy more widely. Banks create money and purchasing power in the economy (McLeay, Radia, and Thomas 2014), and we would expect them to continue to create the marginal unit of money even with a universally accessible CBDC. Paying different rates to banks and nonbanks would allow monetary policy to influence banks—and therefore credit and money creation—differently to nonbanks. For instance, the spread between the rate paid to banks and nonbanks could be used as a monetary policy tool in itself, the implications of which will be briefly explored in section 4.1.

Determining the stance of monetary policy is not the only reason one might pay interest on CBDC. If CBDC is not being used as an instrument of monetary policy per se, policymakers may still want to pay a variable rate of interest in order to regulate demand for CBDC. In this context, CBDC would require two different rates. The interest rate paid on CBDC held by banks would be used to set the stance of monetary policy (just as the rate on reserves is used today), whilst the interest rate on CBDC held by nonbanks could be used to regulate demand for CBDC. Lowering the "nonbank rate" in times of financial stress may help to limit demand for

CBDC and mitigate risks to financial stability from a "digital bank run" (Barrdear and Kumhof 2016).

Some writers have proposed using a negative rate on CBDC, alongside the complete withdrawal of physical cash, as a way of circumventing the zero or effective lower bound on monetary policy. Kimball and Agarwal (2015), Goodfriend (2016), and Rogoff (2016) all suggest that replacing cash with a CBDC could make it easier to set a negative rate on central bank money and thus alleviate the lower bound on interest rates. However, as recognized by these authors, the removal of cash is not a necessary consequence of CBDC and while one may have a bearing on the other, the two policies should be assessed on their own merits. In this paper we do not assume, nor advocate, the removal of physical central bank money.

Were CBDC to be non-interest bearing, it would be closer in spirit to central bank notes. Fixing the rate on CBDC permanently at zero would mean that it could not be used as a monetary policy tool. But if the objectives of the policymaker were to improve payment efficiencies and financial inclusion, it is not essential that a CBDC pays interest. As noted by Armelius et al. (2018), though, a non-interest-bearing CBDC could actually raise the lower bound for interest rates, because it does not bear the storage costs that currently apply to bank notes. This would worsen the constraint on monetary policy, so any benefits to payments would need to be weighed against this cost.

3.3 Convertibility

A third policy choice concerns whether CBDC is freely convertible for (i) other forms of central bank money, namely reserves and cash, and (ii) commercial bank deposits. It is common practice for central banks to allow free convertibility at par (1:1) between their different types of liability (so that £10 of reserves can always be exchanged for £10 of cash). There are clear incentives for central banks to continue with this approach under a CBDC, in order to ensure monetary stability. In this setup the central bank can set the overall size of the monetary base (for example, through asset purchases), but it cannot directly set the composition of liabilities within it (as this is determined by the preferences of banks and nonbanks to hold different types of central bank liability).[2]

However, a number of authors have suggested breaking with this convention, particularly in the context of CBDC. For instance, Kimball and Agarwal (2015) and Assenmacher and Krogstrup (2018) outline a framework in which a flexible exchange rate can be operated between cash and electronic central bank money in order to impose an effective negative interest rate on cash and overcome the lower bound. But such a system would mean that the economy would be operating with two distinct fiat currencies simultaneously, albeit with a managed exchange rate. We doubt that this would be plausible in practice. For instance, it would raise significant questions around which of the two currencies—physical cash or CBDC—would be the unit of account in the economy. Establishing and securing the unit of account is one of the most fundamental roles of a central bank, and so this risk should not be lightly dismissed.

It is also convention that commercial bank deposits (in the form of demand deposits) and central bank money are freely convertible at par. Kumhof and Noone (2018) suggest that ending this convention, so that deposits could not be freely converted into CBDC at par, would help to eliminate the financial stability risks that CBDC may introduce. In our view, the ability of depositors to exchange commercial bank money for central bank money on demand is fundamental in maintaining the confidence in bank deposits, and many of the activities of the monetary authority (such as lender of last resort, liquidity regulations, and deposit insurance) are geared toward ensuring that this is always possible. In fact, we would consider it a necessary part of establishing a stable monetary framework in which CBDC and bank deposits coexisted and exchanged at parity (i.e., 1:1) (Bjerg 2018).

---

[2]This is a central feature of the par convertibility of central bank liabilities even now: the central bank can choose to inject reserves into the banking system, but if there is a surge in demand for central bank notes, then they must allow the stock of notes to increase while the stock of reserves falls, or else inject additional reserves, expanding the monetary base. The alternative is that the increased demand for notes goes unmet and the price of a note changes relative to the price of reserves, breaking parity between different types of central bank money and undermining monetary stability.

3.4 Technology

More practical design choices relate to the technology used to power a CBDC, in particular whether a CBDC should use distributed ledger technology or more established technologies, such as those currently used for existing central bank real-time gross settlement systems. Importantly, this highlights a distinction between a central bank digital currency, which can be based on a variety of technological options, and a central bank *crypto*currency, which by definition is based on cryptographic technology, such as distributed ledgers. While Danezis and Meiklejohn (2016) advocate the use of distributed ledger technology, Scorer (2017) argues that this would not be necessary, even though distributed ledgers may have advantages over centralized ledgers in specific areas—for example, around resilience.[3] However, the technology is still considered to be too immature to power a critical national payment system such as the Bank of England's Real-Time Gross Settlement system (Bank of England 2020).

3.5 Our Benchmark Model of CBDC

Our primary interest in this paper is a CBDC in the context of monetary policy. For most of the analysis that follows we will therefore focus on a particular design of CBDC intended to be used as part of the monetary policy toolkit. Setting the parameters discussed above with this in mind, we consider a CBDC that is universally accessible, interest bearing, and freely convertible to other forms of central bank money and to commercial bank deposits.

We assume free convertibility, at par, for the reasons discussed in section 3.3: without convertibility between its own liabilities, the central bank would be managing multiple distinct fiat currencies, and without convertibility between central bank money and deposits, the central bank risks undermining confidence in the value of those deposits. A consequence of free convertibility between CBDC and deposits is that the impact of a customer withdrawing CBDC on the banking sector's balance sheet is identical to that of a cash

[3]Benos, Garratt, and Gurrola-Perez (2017) evaluate distributed ledger technology in relation to securities settlement and suggest that it has the potential to improve efficiency and lower the costs of trading in these markets, but also highlight the current immaturity of the technology.

withdrawal. When the depositor requests to withdraw CBDC from their deposit account, the commercial bank would reduce the balance of the depositor's account and transfer CBDC across the central bank's balance sheet to the depositor's CBDC account. The banking sector's balance sheet shrinks by the amount of the withdrawal (since reserves/CBDC are removed from its assets while deposits are deducted from its liabilities), whilst the depositor has simply swapped one asset (deposits) for another (CBDC), with no overall change in the size of their balance sheet. Nonbanks are able to convert (freely and at par) between bank deposits and cash (via a cash withdrawal) or between bank deposits and CBDC (via a withdrawal to a CBDC account). Banks are able to convert freely between CBDC (fulfilling the role played by reserves) and cash.

Banks can also convert bonds and other eligible collateral into CBDC, either by selling those assets to the central bank (outright purchases) or using them as collateral to borrow CBDC (through repos). In theory the central bank could allow any economic agent to undertake such transactions and change the aggregate supply of CBDC. However, in practice, this would require central banks to deal with a much wider range of counterparties, with the risk and operational burden that that implies. We therefore assume that for practical reasons, only banks could trigger the creation of new CBDC (in aggregate) by exchanging assets with the central bank. Section 6 discusses the exception to this rule, considering the scenario in which nonbanks are allowed to sell assets directly to the central bank in exchange for newly issued CBDC through programs in the style of quantitative easing.[4]

We assume universal accessibility to take our position to its extreme. As stated already, a limited-access version of CBDC exists already in the form of reserves. Therefore, by comparing the current environment with one with universal access, we can get some sense of an upper bound to the various effects we will discuss below. It is worth being explicit here that we do not assume that reserves coexist alongside a new CBDC. Rather, we assume there is a single, universally accessible CBDC that is available to banks, households,

[4]Again, it should be noted that a number of central banks are currently reviewing who has access to their balance sheet.

and corporates equally, thus replacing reserves as an asset for banks and acting as a new asset for other sectors of the economy.[5]

As our focus is on monetary policy implementation, we assume the monetary authority wants to use CBDC to guide monetary conditions. Therefore, our benchmark CBDC is interest bearing. More specifically, we assume that the central bank uses the interest rate on CBDC as the primary instrument by which to influence other rates in the economy and fulfill its objective to stabilize the macroeconomy. This setup allows us to explore the monetary policy implications when all sectors of the economy can hold CBDC and can respond to changes in interest rates by switching from CBDC to deposits or vice versa. For the most part, we assume that a single rate is paid on all forms of CBDC (i.e., to both banks and nonbanks) and are explicit where we do not.

Beyond assuming that CBDC is practically feasible, we have chosen to take a technology-agnostic approach for the model of CBDC used in this paper. However, even abstracting away from technological choices, some decisions on practical matters are necessary to understand whether CBDC would be widely used. In our benchmark model we assume that the central bank would provide the underlying payments platform for CBDC but would not deal directly with members of the public. Instead, it would allow private-sector firms, such as financial technology firms, payment institutions, or banks, to identify customers, register CBDC accounts on their behalf, and provide the customer interface and customer services related to CBDC accounts. These firms would be responsible for administering CBDC accounts but would not take custody of the CBDC itself, ensuring that CBDC remained a liability of the central bank to the end user, rather than to an intermediary. We assume that the payment services available to CBDC account holders would be comparable to those available to holders of bank deposits, and that the CBDC and deposit-based payment systems would be interoperable (so that any deposit account could make a payment to any CBDC account and vice versa). This means that CBDC would serve as a close—but not perfect—substitute for bank deposits. This point is crucial: as

[5] Making a CBDC universally accessible to nonresidents as well as residents could have implications on capital flows and the exchange rate. These are complex issues that we do not address in this paper.

discussed by Broadbent (2016), if CBDC only serves as a substitute for central bank notes, then the composition of central bank liabilities and private-sector assets changes, but beyond some minor implications for seigniorage, the monetary policy implications are negligible.[6] However, once CBDC starts to offer services similar to bank deposits, there will be an impact on the quantity and price of bank funding, with more interesting implications for monetary policy. Importantly, we assume that CBDC accounts would not offer credit facilities, such as overdrafts, to the vast majority of users, although the central bank would still be able to lend CBDC to certain financial institutions, as it does with reserves today.

4. Controlling the Interest Rate on Central Bank Money

The core of modern monetary policy is the ability to set and guide interest rates in the market(s) for central bank money. Traditionally this has been done in one of two ways, shown in figure 1. In a corridor (or channel) system the central bank sets the quantity of reserves such that the secondary (interbank) market for reserves clears at the target rate. Alternatively, in a floor system the central bank expands the supply of reserves to at least satiate demand. To pin down this demand, the central bank pays a rate of interest on balances, or at least a proportion of balances, held with it overnight (Goodfriend 2002).

Conceptually, very little would change to either of these frameworks if access to reserves were expanded to a universally accessible, interest-bearing CBDC (see figure 2). In a corridor system the central bank could still control the supply of central bank money such that the short-term unsecured borrowing rate for central bank money cleared at the target level. Equally, in a floor system the central bank could expand the quantity of CBDC such that demand was satiated, with CBDC balances being remunerated at a given rate.

One notable difference, however, would be that the market for electronic central bank money would now have a much wider range

[6] The impact on seigniorage could be positive or negative depending on (i) the interest rate paid on CBDC (if any) and (ii) the relative cost of producing CBDC versus the costs of producing the equivalent amount of cash (including fixed and marginal costs). We discuss this more in section 4.

Figure 1. Secondary Market for Central Bank Money without a CBDC

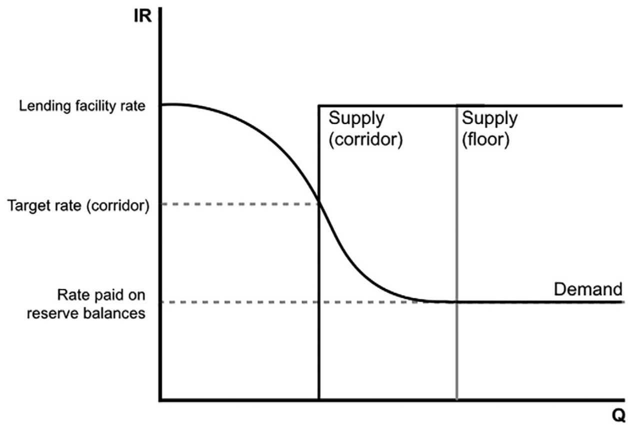
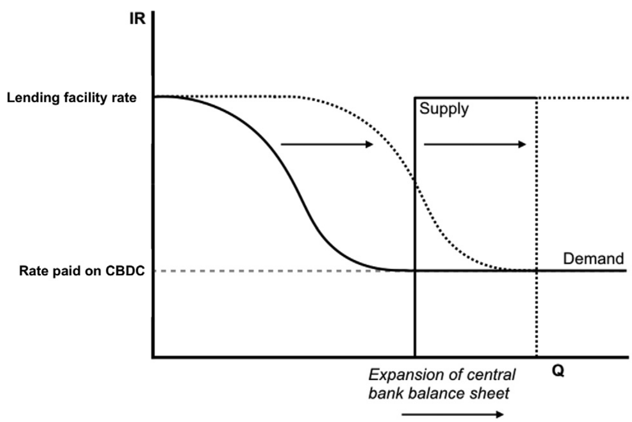


Figure 2. Secondary Market for Central Bank Money with a CBDC



of potential participants, since CBDC would allow electronic central bank money to be held by all agents rather than just commercial banks. These participants may also have a much wider range of motivations for holding, lending, and borrowing CBDC.

A first consequence of this is that the demand curve for electronic central bank money would shift to the right. How far it would shift would depend on the setting of the parameters discussed in section 3 and the extent to which CBDC is a substitute for bank deposits. Assuming that CBDC is freely convertible for deposits, and is integrated into the existing deposit-based payment systems, significant demand for CBDC could come from those who currently hold deposits. This demand could be limited, in the case of households, by the fact that their deposits are covered by deposit insurance, so that there is no significant difference in credit risk between deposits and CBDC. In contrast, for private-sector and nonbank financial corporations, which are not covered by deposit insurance, CBDC offers an asset that has the same payment characteristics as deposits but without any credit risk, and so may be more attractive. Both of these sources of demand for CBDC are likely to be sensitive to the spread between rates paid on CBDC and rates paid on deposits.

Under a corridor system, any change in demand for central bank money would need to be accommodated to ensure that the secondary market continued to clear at the market rate. All else equal, this would imply that the size of the central bank balance sheet would be larger in light of increased demand.

Under a floor system, to maintain the interest rate on CBDC as the relevant policy rate, the central bank would only need to expand supply to the point where aggregate demand continued to be satiated. However, in economies that have implemented large quantitative easing programs, the total expansion in the balance sheet may be less than the total new demand for CBDC from the nonbank sector. QE often results in the banking sector (in aggregate) holding more reserves that it desires, because—prior to the introduction of CBDC—there is no other sector which can hold reserves, and so no way for the banking sector to reduce its aggregate holdings. The introduction of CBDC, which can be held by both banks and nonbanks, would enable banks to divest themselves of unwanted CBDC balances—for example, by buying assets from the nonbank sector in exchange for CBDC. This would allow banks to reduce their holdings of CBDC to the point of inflection in the original demand curve, in the process meeting some of the new demand for CBDC from the nonbank sector. The total increase in the central bank's balance sheet will therefore be equivalent to the new demand for CBDC from the nonbank sector, minus any undesired reserves/CBDC that the banking sector chooses to sell to the nonbank sector (and minus any substitution by the nonbank sector from physical cash into CBDC).

A second consequence of a universally accessible CBDC is that the demand curve would now be subject to a wider range of shocks. The factors that would drive demand for CBDC from corporates, nonbank financial institutions, and households may be different than those that affect the demand of banks. If these new factors and the shocks moving them were positively correlated with those that already affect the traditional market participants, then this, combined with larger absolute balances, would make the demand curve for CBDC less stable than the current demand curve for reserves. In turn, this would mean that a corridor system would require more active management and monitoring by central banks and could potentially lead to a more volatile balance sheet size. Conversely, if the shocks affecting new market participants were negatively correlated with those of existing participants, the demand curve could actually become more stable, as there would now be a deeper market that could smooth the idiosyncratic deviations of different participants. What is more, even if the demand from new participants was volatile, that would not necessarily mean it was less predictable than currently. Demand for cash is subject to large seasonal swings in demand, but central banks are able to manage these relatively easily, as they are often predictable and based around weekends, holidays, or similar shifts in spending. Whether the demand curve for CBDC would be more or less difficult to manage is ultimately an empirical question, though one that is hard to answer ex ante. Its practical importance, however, makes it a high priority for future research.

Universal-access CBDC would also have implications for the extent to which the interest rate paid on central bank money could act as an effective floor to other interest rates. A number of central banks employing floor systems have observed market rates falling below the interest rate paid on overnight balances at the central bank. One of the main causes of this is the tiered access to the central bank's balance sheet in these economies. In a floor system, when only a limited number of agents have access to the central bank's balance sheet and are paid interest on their balances held there, those agents who do not have access or do not receive interest are willing to lend funds overnight for less than the rate paid by the central bank, making the floor less effective (Keating and Machiavelli 2018). Giving everyone access to the central bank's balance sheet via a universally accessible CBDC and paying all market participants a rate of interest would better ensure that no one is willing to lend for a rate lower than the floor, making the floor a more effective constraint.

The Federal Reserve's Overnight Reverse Repo Program (ONRRP) provides a useful test case for the potential benefits of a universally accessible CBDC. In 2013 the Federal Reserve widened access to central bank liabilities by allowing a range of nonbank money market investors to participate in the ONRRP. Prior to this, short-term rates in money markets had fallen below the interest on excess reserves (IOER) paid to bank counterparties. Since the introduction of the ONRRP, the Federal Reserve has been able to impose a firm floor on short-term interest rates. Although it represents a more limited expansion of counterparties than the case of universal access that we explore in this paper, the Fed's experience with ONRRP highlights many of the design tradeoffs required to balance the monetary policy and financial stability objectives that are relevant to our analysis, including the potential to alleviate some of these with differentiated rates, as discussed by Frost et al. (2015).

One question that widening access to electronic central bank money raises is: under a corridor system, which rate in the secondary market should the central bank target? Currently, participants in the market for reserves lend to one another on relatively short terms—almost exclusively less than 12 months—as their main motivation is to smooth temporary liquidity shocks. However, assuming that nonbanks engage directly in the market, a more diverse set of market participants may lead to CBDC being lent for a wider range of reasons and over a wider set of terms. It is conceivable that a corporate looking to borrow for capital expenditure may take out a CBDC-denominated loan over a number of years, or that a household may borrow CBDC to buy a house, in which case the terms could be measured in years rather than days or months. However, as long as the sellers of goods, services, or assets are able to substitute freely between deposits and CBDC (which we assume they can), then they should be neutral between receiving payments in CBDC or deposits, and should not offer any incentive for buyers to pay by one medium over the other. This means borrowers themselves should be neutral between borrowing CBDC and borrowing deposits—they will borrow from whichever lender offers them the best borrowing rate.

This would create a term structure of central bank money in a way that does not currently exist. The shape of this term structure would depend on expectations of the overnight rate for CBDC. The interest rate on a given loan would then be a function of this expected path and a series of premiums based on factors including the length of the term and the risk of default.

Even with a larger term structure of rates for central bank money, it makes sense for policymakers to continue to target very short-term rates, as they largely do now. In part this is because these rates contain fewer premiums. By effectively controlling short rates, and giving clear forward guidance about the future path these rates will take, policymakers can also guide longer rates. However, an additional reason is that in the case of a universal CBDC, it is likely that the participants in the short-term segment of the market would not change very much. Households and nonfinancial businesses rarely manage their liquidity at the overnight frequency and so would not likely borrow at such a short term. Conversely, if they were willing to lend to a bank overnight, then they would likely already be holding their funds in the form of demand deposits. Some nonbank financial corporates may be more willing to be involved in the overnight market, but this is likely to be skewed toward those firms that already have a working knowledge of markets that are very similar to that for overnight central bank money. To the extent that this happens, it should make the overnight market both more competitive and more liquid. Crucially though, the large majority of overnight lending will still be undertaken by the banking system and firms who already have access to electronic central bank money, allowing central bankers to guide the market much as they do today.

On balance, none of the above would appear to restrict the central bank's ability to control the short-term interest rate on central bank money, although a more volatile demand curve might make a floor system more attractive as an operational framework than a corridor. Widening access to central bank money could even afford policymakers tighter control of rates when operating under a floor system, because the risk-free rate on central bank money is available to all agents rather than only the banking sector.

4.1 Paying Different Rates to Banks and Nonbanks

Although many central banks opt to pay a single rate of interest on balances deposited with them, this need not be the case. A number of central banks now pay different rates of interest on balances based on the amount held, so as to influence the interest on the marginal unit differently to that on the entire balance (see, for instance, Basten and Mariathasan 2018 for a discussion of the Swiss case). This would be perfectly viable with a universally accessible CBDC, and could allow the central bank to influence those agents who use large holdings of CBDC as a store of wealth differently than agents who hold small balances purely for transactional purposes. This could be useful if policymakers wanted to impose negative rates on large, idle holdings of CBDC without penalizing smaller transactional accounts.

Similarly, a central bank could pay different interest rates depending on who holds the CBDC.[7] Perhaps the most logical way to differentiate CBDC holdings with a universally accessible CBDC would be between balances held by banks and those held by the nonbank private sector. This could be motivated by the special role that banks play in the monetary transmission mechanism and the economy more widely. Banks create money and purchasing power in the economy (McLeay, Radia, and Thomas 2014) and we would expect them to continue to create the marginal unit of money even with a universally accessible CBDC (see section 5.3). Paying differentiated rates of interest would allow monetary policy to influence banks—and therefore credit and money creation—differently to nonbanks.

The spread between the rates paid to banks and nonbanks could be set as a positive, but fixed, level with the two rates moving together. In this case there would be a steady-state impact on banks' balance sheets, but the consequences for monetary policy beyond our benchmark case would be limited. Alternatively, the spread itself could be varied through time, which would have implications for monetary policy. An increase in the spread could lower the cost of holding a given level of liquidity for a bank, which in turn would improve its balance sheet position. There would therefore be a

[7]For instance, not all participants in the federal funds market currently receive interest on their balances held with the Federal Reserve.

monetary stimulus through the bank balance sheet channel, or through the bank lending channel. Even if banks passed some or all of the increase in the return on their assets to depositors in the form of higher deposit rates, this would be stimulative through a cash flow channel.[8] The spread between the banking-sector CBDC rate and the nonbank CBDC rate would therefore provide an extra dimension to the stance of monetary policy. Such a design feature may also be attractive to a policymaker introducing CBDC for financial stability motives, as it would allow them to incentivize the holding of liquidity in specific sectors of the economy.

Although paying different rates to banks and nonbanks may assist in implementing monetary policy, a system that explicitly pays more to banks than to the public at large could be politically contentious; these debates may be especially acute in times of financial crisis.

4.2 Seigniorage

In section 3.5 we highlight that the impact of a universally accessible CBDC would have an ambiguous impact on seigniorage. While the net impact on the extent of seigniorage would be uncertain, it may also be relatively small in terms of overall monetary transmission. However, a more fundamental issue would be if the changes were such that seigniorage actually became negative, with an increasingly quasi-fiscal dimension to monetary policy.[9] While theoretically possible, the conditions required for such an eventuality are limited and seem unlikely. For instance, positive seigniorage is practically guaranteed if the central bank were to control the aggregate supply of CBDC through outright purchases. With the rate paid on CBDC as the floor of rates in the economy, the assets held on the central bank's balance sheet would always imply a greater return and positive seigniorage.[10]

---

[8] Banks may choose to pass on the higher return in order to strategically gain market share, or because if they didn't, another bank would.

[9] We thank an anonymous referee for highlighting this point for us.

[10] Obviously the central bank would be subject to interest rate risk and could ultimately make losses if it failed to manage the risk on its balance sheet correctly, but this is an issue with which central banks are familiar, and on which they have significant expertise.

Were the central bank to supply CBDC via repo transactions, then both the rate of return paid on CBDC balances and the rate with which it conducted repos would be decision variables of the policymaker. For negative seigniorage to occur, the central bank would have to pay a higher rate of interest on balances than it charged to issue new CBDC through repos. This seems unlikely, as it would amount to providing CBDC at below cost. It would also undermine the use of the interest rate paid on CBDC as the policy rate, as it would open up a risk-free arbitrage opportunity to borrow CBDC from the central bank via a repo and then simply store it in a CBDC account. This would lead to the floor on market interest rates being bounded on the downside by the repo rate, not the rate of interest paid on CBDC. Were the central bank to pay differentiated rates on CBDC and pay nonbanks a higher rate on balances than it paid banks, or charged banks to borrow through repos, then there could potentially be negative seigniorage across the entire balance of CBDC. However, as argued above, the repo rate and interest on CBDC rates are both decision variables of policymakers and it is more likely that differentiated rates would be used to pay higher interest rates to banks than to nonbanks, if at all.

4.3 Issuing CBDC Alongside Existing Reserves

An alternative to paying different interest rates on a single form of electronic central bank liability would be for the central bank to issue CBDC as a distinct asset alongside existing reserves. The rate of interest paid would depend on the asset held and not the counterparty holding the asset. This is similar to the situation today: physical cash is distinct from reserves, and pays zero interest regardless of whether it is held by a bank or a nonbank. Nonbanks can only hold cash, whereas banks can hold both cash and reserves, meaning that they have access to a preferential rate of interest (i.e., the policy rate paid on reserves). This arrangement has parallels to the model assumed in Barrdear and Kumhof (2016): banks would be able to hold both reserves and CBDC as distinct assets bearing different interest rates, whilst nonbanks would only be able to hold CBDC. The primary rate of monetary policy would be the rate paid on reserves, while the rate paid on CBDC would be used to control demand for CBDC relative to bank deposits.

At first glance this scenario seems to offer extra possibilities for monetary policy under a CBDC. However, under some sensible assumptions, we believe the outcome would be practically identical to the scenario above in which different rates are paid to banks and nonbanks on a single liability in the form of CBDC. As we show in Meaning et al. (2018), if different forms of central bank liabilities are freely convertible, and reserves and CBDC are equivalent in a technological, transactional, and regulatory sense, banks will choose to hold only reserves as long as they pay a higher rate than CBDC. The end result is that banks are paid the reserves rate and nonbanks earn the CBDC rate, equivalent to the one-asset framework outlined above.

For the rest of the analysis we assume that there is only one form of CBDC paying one rate regardless of whether it is held by banks or nonbanks.

5. Further Transmission of Monetary Policy

5.1 Transmission to the Wider Universe of Interest Rates

As is currently the case, the process of arbitrage would mean that changes in the rates on CBDC would pass through to rates on other assets in the economy. In the appendix we present a stylized model that provides an intuitive formalization of this process under a universally accessible CBDC. In essence what one would expect to see is that the interest rate that clears the market for CBDC would be the theoretical risk-free rate, minus a premium derived from the nonpecuniary transactional utility (or convenience yield) provided by CBDC:[11]

$$R^C = R - \phi^C. \tag{1}$$

This transactional utility can be motivated in a number of ways and need not be constant, but it could vary through time and, as posited by Friedman (1960), it could be a function of the stock of CBDC

[11] The market rate may also be a positive function of the cost of administering accounts, or the underlying payments system. ϕ^C could therefore be thought of as a more complex composite premium of factors that drive a wedge between the risk-free rate and the rate on CBDC.

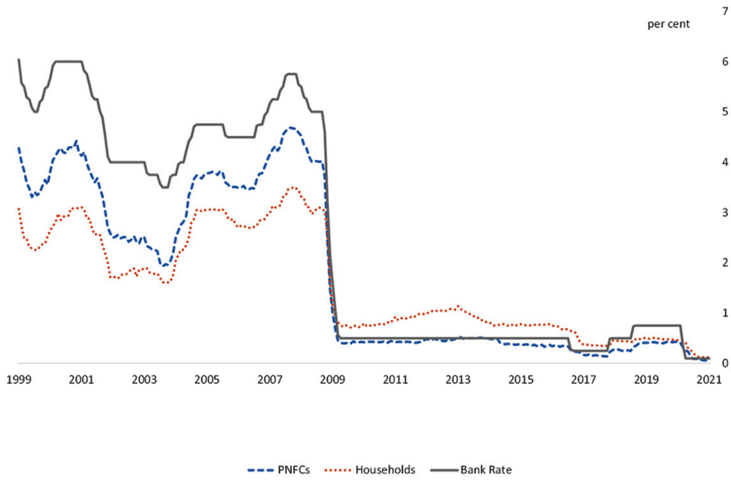
itself. Bordo and Levin (2017) develop the ideas of Friedman in the context of CBDC. They suggest that if $\phi^c = f(Q^c)$ and the creation of CBDC is costless, then the supply of CBDC should be expanded to the point where $\phi^c = 0$ and the rate of return on CBDC equals the risk-free rate, $R = R^c$. The rates of interest on other assets in the economy can then be written in the general form by

$$R^x = R^c + \phi^x, \tag{2}$$

where ϕ^x is a premium specific to asset x that could be a function of a wide range of factors. These factors could include the relative probability of default compared with CBDC, the relative transactional utility provided by the asset, and the relative utility provided by the asset for liquidity-management or regulatory purposes. For instance, government bonds would have little to no default risk, so that element of the premium would be low, but they also provide little to no transactional utility, which would cause a positive spread against the CBDC rate to occur. Deposits with commercial banks could provide transactional services similar to CBDC, and so would see an equivalent premium along that dimension, but inherently contain some risk of default, and so would see a positive spread over the CBDC rate.

An initial consequence of a universally accessible CBDC is likely to be that deposit rates offered by commercial banks would most likely need to be higher than the interest rate paid on CBDC. Prior to the 2008 financial crisis, deposit rates were consistently below the policy rate (figure 3). This could be explained by two phenomena. First, banks had a monopoly on the creation of *electronic* money that could be used to make transactions in the wider economy and nonbanks had no outside alternative. Banks could make use of this monopoly to lower the rate that they paid on deposits. In terms of equation (2) this meant that the transactional utility for nonbanks of reserves was zero, while the transactional utility of commercial bank deposits was high, so the latter rate naturally fell below the former. What is more, state backstops, implicit or otherwise, gave the perception that bank deposits were essentially risk free, so there was no default premium built into deposit rates. Second, banks offered a bundling of services, such as overdrafts and preferential mortgage or lending rates, which provided a utility to holding deposits with

Figure 3. Sight Deposit Rates vs. Policy Rates in the United Kingdom



Source: Bank of England.

the banking sector and which were reflected in a negative premium, lowering the rate at which the deposit market cleared.

With a universally accessible CBDC, central banks would be making available to nonbanks an outside, competitive option for the provision of a means of payment (although unlike some bank deposits, this would not be bundled with other services such as overdrafts). With such an outside option, depositors who are offered less than the current CBDC rate have the option of withdrawing CBDC (in exchange for a reduction in their deposit balances). This would, as a minimum, lower the spread between the policy rate and deposit rates. If the additional services offered by deposits held with a commercial bank, such as overdraft facilities, were not sufficient to outweigh the additional risk they represented, then the spread would, in fact, change sign and deposit rates would need to be higher than the CBDC rate. This would constitute a large structural change for the banking sector which could have consequences for credit provision and banks' funding models. Some of these effects, such as pressures on net interest margins, could be largely if not completely mitigated by the paying of differentiated rates of interest on CBDC balances, as discussed in section 4.1.

As well as potentially changing the structure of interest rates in the economy in a levels sense, there are a number of features of the model of CBDC outlined earlier that could affect the speed and extent of pass-through between changes in the policy rate and other rates. Most notably, the existence of a competitive money alternative to bank deposits is likely to mean that if the interest rate on that alternative changes, but deposit rates do not move by an equal amount, then people will reallocate their portfolio to take advantage of the relatively higher return that has opened up. This would create flows between the two assets. If the policy rate which is paid on CBDC is increased, then this could result in a fall in demand for bank deposits, while if the policy rate is cut, this could drive demand from CBDC into bank deposits. This will be particularly acute when it is easier to convert between CBDC and deposits. To the extent that pass-through from policy rates to deposit and wholesale rates has been estimated to be currently less than one, CBDC is likely to strengthen this stage of transmission.

We agree with Bordo and Levin (2017) that deposit-taking institutions that engage in customer-focused relationship banking are likely to be less vulnerable to the outflows of deposits than the areas of the market that compete purely on price terms. We also acknowledge that evidence for the United Kingdom shows that deposits are sticky—depositors do not often switch banks—and so have low price elasticity (Chiu and Hill 2015). However, deposits may become less sticky once the central bank offers an outside option, and as a result of regulatory changes that will make it easier to switch accounts (such as the European Union's Second Payment Services Directive). Ultimately, whether there is a change in the speed of pass-through from changes in the policy rate to changes in deposit rates will depend on how banks react. We could assume that technology is likely to mean that depositors can more easily and costlessly move between deposits and CBDC, so banks would have to respond quickly to stem deposit flows. However, banks may react to the increased flightiness of deposits by changing their funding models to rely more on term funding, to "lock in" deposits. This would mean that pass-through to rates paid on deposits would ultimately be slower.

Larger changes in deposits and wholesale rates for a given change in the policy rate would also mean a larger impact on banks' funding costs, all else equal. For a given markup on these funding costs, this would have a larger impact on loan rates. This increased sensitivity of both funding costs and lending rates to changes in the policy rate could act to strengthen the bank *lending* channel, which we will discuss further in the next section.

5.2 Transmission to the Real Economy

The ability to influence interest rates is only an intermediate step in the monetary transmission mechanism. We now consider the implications of a universally accessible CBDC for the transmission of policy changes to the real economy.

Given that our previous analysis suggests that universally accessible CBDC is likely to make interest rates more sensitive to changes in the policy rate, all else equal, this will serve to amplify the strength of a number of channels for a given change in the policy rate. Most obviously, the real interest rate channel would strengthen as interest rates on savings and credit would shift more for the same change in the policy rate and so would the incentives for intertemporal substitution by economic agents. For similar reasons, the cash flow channel would also be strengthened.

The impact on the bank lending channel of transmission is less clear. To the extent that bank funding costs would become more sensitive to changes in the policy rate for the reasons described above, this would strengthen the bank lending channel. What is more, the additional competition in credit provision may make pass-through to lending rates more complete. However, there are a number of factors that would most likely weigh in the other direction. First, were deposit rates to be above policy rates, this could squeeze the net interest margins of the banking sector, which could result in lower profits. This would mean bank capital grows at a slower rate, constraining banks' ability to lend and therefore weakening the bank lending channel.

Perhaps more fundamentally, if a CBDC were to disintermediate the banking sector and significantly reduce the size of its aggregate balance sheet, this would reduce the importance and traction of the bank lending channel. This disintermediation is more likely when a

CBDC is a close substitute for bank deposits and fully or partially dominates them in some aspects. This would be precisely when the benefits to the monetary transmission mechanism discussed above would be at their greatest. This serves to highlight an important tradeoff in the setting of the design parameters of CBDC: making CBDC too attractive relative to bank deposits will strengthen the real interest rate channel by providing a more competitive outside option, but weaken the bank lending channel by disintermediating banks. Conversely, making CBDC very unattractive relative to deposits means that it would not represent a true outside option to deposits, so very little would change. How to manage this tradeoff will depend on many factors, including the importance of the banking sector in the economy, the viability of nonbank finance to provide credit to the economy in lieu of a diminishing banking sector, and the central bank's ultimate motivation for introducing CBDC. It would seem likely that if the motivation for CBDC is purely to provide a secure digital payment system with no need to affect monetary policy, then it is unnecessary to make CBDC an attractive substitute for *interest-bearing* deposits, especially when weighed against the risk of disintermediating the banking sector.

5.3 CBDC and Money Creation by Banks

The transmission mechanism may also change if the introduction of CBDC affects the way that banks choose to issue loans. Banks are crucially involved in money creation in the economy, because they lend by issuing new deposits, in effect creating new money and purchasing power (McLeay, Radia, and Thomas 2014; Jakab and Kumhof 2015). In contrast, nonbank lenders transfer existing purchasing power (either deposits or CBDC) from savers to borrowers, but do not create any new purchasing power in the process. With the introduction of CBDC, banks could still continue to lend by issuing new deposits, but they would now have the option to lend CBDC (by transferring CBDC to the borrower's CBDC account).This would make them more like nonbank lenders, and reduce the sensitivity of money creation to a change in monetary policy.

In practice, however, there are a number of reasons why banks would continue to prefer to lend by issuing new deposits. Firstly, as discussed previously, there is unlikely to be demand for borrowing in CBDC specifically, implying that the interest rate on loans for a given level of risk and term should be the same whether it is CBDC or deposits that are borrowed. Secondly, for a bank, lending CBDC will have a more negative impact on current regulatory ratios (specifically the liquidity coverage ratio) than lending via issuing deposits. Lending CBDC ensures that the bank will lose £100 of liquidity for every £100 lent. In contrast, while lending by issuing deposits could still lead to some outflow of CBDC, it is likely to be less, on average, than 100 percent of the amount lent. Given this, credit creation by the banking system would continue as now even with the existence of a universally accessible CBDC, with new loans initially matched by newly issued bank deposits, and money creation would continue to be sensitive to changes in monetary policy. Of course, once those new deposits are created, though, it would remain the optimal portfolio choice of nonbanks as to how much of the newly created bank deposits were converted to CBDC, as is now the case with regards to the substitution between deposits and central bank notes.

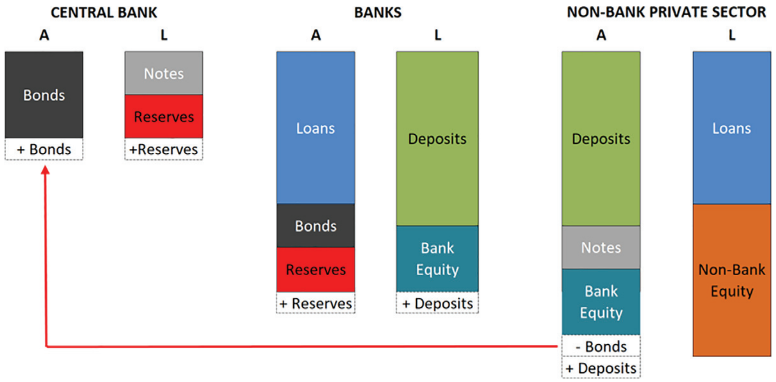
5.4 Overall Impact on Transmission

Overall the analysis above suggests that a universally accessible CBDC would most likely strengthen the impact of changes in the policy rate on the real economy, predominantly through increased pass-through from policy rates to other interest rates in the economy. Importantly, to the extent that this were true in practice, a CBDC would imply that policy rates needed to vary by less over the cycle to stabilize the economy, conditioned on the same shocks afflicting the economy. This could have benefits to policymakers, especially in the context of the lower bound as, for a given steady-state level of policy rates and set of shocks hitting the economy, the probability of the policy rate becoming constrained would be reduced.

6. Asset Purchases with CBDC

In recent years central banks have used newly issued money to purchase financial assets from the private sector, a policy known as quantitative easing (QE). This policy aimed to manipulate the relative supplies of assets in the economy and to lower longer-term interest rates, given that short-term rates were constrained by their

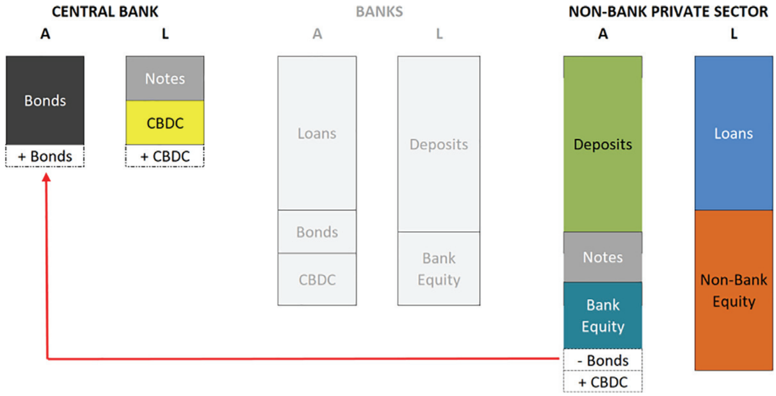
Figure 4. Central Bank Asset Purchase without a CBDC



effective lower bounds. However, since most nonbanks cannot currently hold accounts at the central bank, and it is impractical to pay for these assets by printing central bank notes, central banks had to use commercial banks as intermediaries between them and the ultimate (nonbank) sellers of the assets. In this process, the nonbank holds a deposit account at the commercial bank, which in turn holds a reserves account at the central bank, as shown in figure 4. (For color versions of figures 4 and 5, see the online version of the paper at http://www.ijcb.org.) The commercial bank sells the asset to the central bank on behalf of the ultimate seller (the nonbank). The central bank pays for this asset by crediting the commercial bank's account with newly issued reserves. Simultaneously, the commercial bank credits the nonbank's account with a new deposit. In this way, asset purchases through QE expand both narrow money (directly) and broad money (indirectly) by the same amount, through the simultaneous creation of new reserves and deposits.

If CBDC is introduced, the central bank would no longer need to use commercial banks as intermediaries to *indirectly* purchase assets from the private sector. Instead, the central bank could purchase financial assets *directly* from the nonbanks, paying for these assets by crediting the nonbank's CBDC accounts with newly issued CBDC, as shown in figure 5. This process does not alter the balance sheets of commercial banks and has no impact on the amount of central bank money held by the banking sector. This means that

Figure 5. Central Bank Asset Purchase with a CBDC



the effect of the policy would not be dependent on the reaction of the banking sector to holding increased quantities of reserves.

When asset purchases are conducted *indirectly* (i.e., through the banking system, prior to the introduction of CBDC), this reaction can take two forms: there is a mechanism through which QE may expand bank lending, making QE more effective, and an opposing mechanism through which banks' reactions may directly offset some of the desired effect of QE on interest rates.

With regards to the first mechanism, the impact of QE on the bank lending channel is presented in a theoretical model by Christensen and Krogstrup (2019). This states that the exogenous increase in reserves held by banks changes the characteristics of banks' balance sheets, lowering the degree of maturity mismatch and relaxing liquidity constraints away from their pre-QE preferred state. Banks respond by rebalancing their portfolios. In particular, they may increase bank lending, amplifying other channels of QE that work outside of the banking system.

If QE is undertaken directly with the nonbank sector using CBDC, there will be no impact on bank balance sheets, and so no need for banks to rebalance in this way. This would weaken any effects of QE through the bank lending channel. However, empirical work by Butt, Churm, and McMahon (2015) suggests that the impact of QE on the bank lending channel is already small, at least in the context of QE in the United Kingdom, and so the importance of this change in effect may be limited.

The opposing mechanism, by which banks may offset some of the desired effect of QE, depends on the extent to which banks meet their regulatory requirements for high-quality liquid assets (HQLA) by holding both government bonds and central bank money. Assuming that banks hold their preferred mix of government bonds and central bank money prior to QE, the effect of QE is to leave banks holding more HQLA than they desired and a greater-than-desired proportion of that HQLA in the form of reserves rather than bonds. The banking sector has no choice, in aggregate, but to hold this larger stock of reserves. If the sector wants to reduce its stock of HQLA to its pre-QE level, then it must sell off an amount of bonds to the nonbank sector. Simple rolling correlations on U.K. data suggest that as QE increased the quantity of reserves in the system, banks reduced their holdings of government securities in a manner consistent with them substituting from bonds to reserves. This sale of government bonds by the banking sector would push down prices and push up yields at exactly the time that the central bank is aiming to lower bond yields by buying them in large quantities, acting as an offset to the desired effect of QE.

This is different from the situation in which a universally accessible CBDC enables central banks to conduct purchases directly with nonbanks. To see clearly why, consider the balance sheet and decisions of the nonbank seller in the previous example. From their perspective, they have exchanged a government bond for a digital form of money. In a world of universally accessible CBDC, the exchange is largely the same. Given this, one could reasonably expect them to rebalance their portfolio in a broadly similar way and drive similar price responses, all else equal.[12] The difference between the two examples is therefore centered on the banking-sector portfolio and banks' demand for government bonds and reserves compared with one another and other assets. Following the arguments above, banks' demand for government bonds following a QE operation would be greater in the case of a universally accessible CBDC. This would result in lower rates on government bonds following asset

[12]To the extent that CBDC is more risk free than a bank deposit, exchanging it for a risky asset may actually represent a larger change in portfolio characteristics and so induce more rebalancing or larger price movements.

purchases.[13] Put another way, the current structure of QE programs creates additional money both on the nonbank and banking-sector balance sheets. QE with a universally accessible CBDC would only create additional money in one of those places.

Other key elements of QE would remain unchanged. For instance, in aggregate, the system would be unable to rid itself of CBDC. The closed system would now be extended to include more than just banks, so banks could shift demand to nonbanks, and vice versa, but importantly, an increase in quantity would have to be absorbed somewhere, and thus prices of other assets would have to adjust to allow portfolios and markets to clear.

Consequently, QE conducted directly with the nonbank sector through the use of a universally accessible CBDC would be stronger insofar as it would not directly affect banks' demand for government bonds as HQLA, but this would come at the cost of weakening the bank lending channel of QE. The net result is unclear. On balance, the fact that the central bank is able to directly engage with a larger pool of direct counterparts and can avoid distortions to the banking sector's aggregate balance sheet leads us to believe the benefits would outweigh the negatives, increasing the overall efficacy of QE. It should also be noted that a universal CBDC does not preclude the central bank conducting asset purchases directly with the banking sector, should it choose to.

Going further, some authors suggest that a universally accessible CBDC would make quantitative easing as a policy instrument entirely obsolete. Bordo and Levin (2019) argue that abolishing physical cash and replacing it with an interest-bearing CBDC would remove the lower bound on nominal interest rates and allow central banks to conduct policy solely through altering the interest rate on CBDC. In a theoretical economy where cash does not exist, this is certainly true. However, although usage of cash is declining in Sweden and a few other developed economies, it remains

[13]Of course, the relative price changes between bonds and other assets that result from QE would likely lead to a second-round endogenous response from the banking sector. This could also reduce demand for government bonds, because the higher price and lower return would make them relatively less attractive compared with central bank money or other assets that could contribute to their HQLA. However, this mechanism would exist whether QE was conducted via issuance of traditional reserves or a universal CBDC.

significant in the majority of countries, including in the United Kingdom, Germany, and Japan. Therefore it is likely that, in most countries, CBDC would coexist alongside physical cash at least for the foreseeable future and would be a useful tool to enhance the effectiveness of QE.

7. Risks to Financial Stability and the Central Bank

Beyond its implications for monetary policy, CBDC could have implications for financial stability. Broadbent (2016) and Carney (2018) among others have highlighted the potential risks of large-scale "digital bank runs." A detailed analysis of these issues is a topic for another paper, but it is worth noting that measures to prevent this kind of risk may have consequences for the monetary policy impacts of CBDC. For example, introducing frictions that discourage large-scale runs to cash, such as daily transfer limits, notice periods for large CBDC withdrawals, or imposing fees on unusually large balances that could approximate the storage costs of cash, would affect the attractiveness of CBDC. This would reduce the extent to which it was a substitute for bank deposits and in turn limit the potential benefits from CBDC acting as an outside competitive option.

CBDC could also impose risks on the central bank, as a result of the need to hold assets to back a larger central bank balance sheet. Were the central bank to look to back these with incredibly safe assets, such as government bonds, then high demand for CBDC could lead to the central bank holding a significant fraction of domestic government debt. This could be especially acute in moments of financial stress when demand for safe assets increases; demand for both government bonds and CBDC could increase simultaneously. Alternatively, the central bank could back CBDC with a wider range of assets, but this would imply a riskier balance sheet than many hold now. Central banks would need to both manage these risks and be aware of the impact they would be having on any new asset markets in which they were to engage.

Further work is therefore needed to investigate the interactions between the impacts of CBDC on monetary policy, financial stability, and the central bank balance sheet. This work should draw on the growing literature on the interactions of current monetary-macroprudential policies (Nier and Kang 2016).

8. Concluding Remarks

At its simplest, a central bank digital currency can be thought of as electronic narrow money, and so in many ways it should feel familiar to economists and policymakers. Within this general definition there are important policy choices to be made, including around access, convertibility, and whether CBDC is interest bearing. Careful consideration will need to be given to these policy choices in future research and in order for any central bank to effectively introduce CBDC.

The main conclusion of this paper is that under a universally accessible, freely convertible, and interest-bearing CBDC, monetary policy could operate in much the same way as it currently does, guiding the economy through varying the rate of interest on electronic central bank money and the aggregate quantity of that money. The untested nature of such a CBDC means that the impact on the monetary transmission mechanism is uncertain, but we believe the most likely consequence is that CBDC would strengthen the monetary transmission mechanism, for a given change in policy instruments.

As well as enhancing the functioning of the existing monetary toolkit, CBDC also has the potential to enable more significant change in the toolkit itself. This paper has only considered the impact of CBDC on widely used monetary policy, namely changes in the central bank's policy rate, and asset purchases (quantitative easing). However, CBDC has also been discussed in the context of less conventional monetary policy, such as the use of negative rates (Kimball and Agarwal 2015) or direct distributions of newly issued CBDC to citizens—so-called helicopter money (Turner 2015; Hampl and Havranek 2018). These policies do not necessarily require a CBDC to be implemented, but the existence of CBDC may affect their feasibility and impact. It may even be that CBDC and future technological progress give rise to monetary policy instruments that have not yet been considered. This is a promising area for future research.

This paper is intended as an early step in the development of the literature on central bank digital currencies, and many fundamental questions remain unanswered. These relate to, among other things, the impact on the wider financial sector, the implications for financial stability, steady-state changes in the economy resulting from CBDC, and how CBDC would affect the balance sheet management of central banks. Significant work, both theoretical and empirical, will be required to inform any policy decision to introduce CBDC.

Appendix. The Structure of Interest Rates and Arbitrage Conditions

To aid thinking about the structure of interest rates in an economy with a universally accessible CBDC, we present a stylized model of the key arbitrage conditions that might prevail. While this model is partial rather than general equilibrium, and appeals to sensible assumptions rather than strict microfoundations to motivate the range of premiums, it still serves as a useful expositional tool and offers some initial, intuitive insights.

We begin from a theoretical risk-free interest rate, R, that represents the return on a pure store-of-value asset with no associated premiums. There is no risk of default, no illiquidity, and no term. CBDC is assumed to be risk free in the same way, but alongside its store of value function it also provides an additional service as a means of exchange, for instance, lowering transaction costs, ϕ^C. The total expected return from a unit of CBDC is therefore[14]

$$R^C + \phi^C \tag{A.1}$$

and no arbitrage would imply

$$R = R^C + \phi^C, \tag{A.2}$$

meaning that

$$R^C = R - \phi^C. \tag{A.3}$$

[14]With no default yet in the model, the expected returns are equal to the agreed returns.

This means that CBDC would clear at a rate below the theoretical risk-free rate by a spread determined by the transactional utility supplied by CBDC. In the exposition here we assume this transactional utility is fixed per unit of CBDC, independent of the quantity of CBDC. This is purely for clarity of presentation, and a credible alternative assumption is that the degree of transactional utility is a function of the quantity of CBDC held. Were transactional utility to be a negative function of quantity, but to a decreasing extent (the implied function has a negative first derivative and positive second derivative), then Friedman (1960) argues that, as the creation of central bank money is costless, the supply should be expanded to the point where ϕ^C is zero and $R^C = R$. This would imply that the rate paid on CBDC could be considered a reflection of the true risk-free rate.

Building from CBDC to a wider array of assets, we look at each sector of the economy in turn and work through the no-arbitrage conditions that their balance sheets would imply.

Beginning with the nonbank private sector, consistent with our stylized balance sheets (see figures 4 and 5 in main text), they can hold their wealth as a combination of three assets: CBDC (denoted by C), bank deposits (D), and government bonds (B).[15] For simplicity of exposition we assume that each of these assets is one period in term but can be differentiated by other characteristics. As discussed above, each unit of CBDC held provides a nonpecuniary benefit to the holder, ϕ^C—for instance, as a result of lowering transactional costs. Similarly, the interest rate paid on bank deposits is R^D, and bank deposits offer a similar but not necessarily equivalent nonpecuniary return emanating from its role as a means of exchange, ϕ^D. However, unlike CBDC, there is a probability, γ, that banks will default on their deposits, in which case the depositor gets neither the pecuniary return nor the nonpecuniary benefit. Lastly, government bonds offer an interest rate R^B. They are assumed to offer no transactional services but are defaultable with probability δ. Taken all together, this means that the nonbank's end-of-period objective function can be written as

[15] We abstract from equity without any loss of insight for the themes with which we are concerned.

$$R^C C + \phi^C C + (1-\gamma)[R^D + \phi^D]D + (1-\delta)R^B B, \tag{A.4}$$

and so is maximized where

$$\frac{dU}{dC} = R^C + \phi^C = 0 \tag{A.5}$$

$$\frac{dU}{dD} = (1-\gamma)[R^D + \phi^D] = 0 \tag{A.6}$$

$$\frac{dU}{dB} = (1-\delta)R^B = 0. \tag{A.7}$$

For the nonbank sector then, assuming that the rate on CBDC is set by the central bank,

$$R^D = \frac{R^C + \phi^C - (1-\gamma)\phi^D}{(1-\gamma)} \tag{A.8}$$

and

$$R^B = \frac{R^C + \phi^C}{(1-\delta)}. \tag{A.9}$$

This gives rise to two spreads. The spread of the deposit rate over the CBDC rate is a positive function of the relative transactional services of CBDC compared with deposits, and a positive function of the default rate. The spread of bond rates over the CBDC rate is a negative function of the transactional service of CBDC money, and a positive function of the default risk in the government bond. This all occurs in a one-period setting. In practice there is likely to be another significant premium driving a wedge between the two rates, which is the term premium. This could be derived in a multi-period model by the additional risk of locking funds away when you are subject to unknown payments or liquidity shocks, and would appear as a positive function of the term of the bond.[16] As discussed previously, were we to assume that the transactional utility of CBDC were a

[16] As another point of reference, bond rates will differ from the theoretical risk-free rate discussed above to the extent of inherent default risk only. In practice, there would also be other premiums, such as term, which would mean that the bond rate clears at a spread above the risk-free rate.

decreasing and concave function of the quantity of CBDC, then the supply can, and arguably should, be expanded to the point at which $\phi^C = 0$. This would mean that the only differences between government bonds and CBDC were the default risk in government bonds and the term. For short-term government bonds in stable economies, both of these elements could be expected to be negligible, and so we would expect the short-term government bond rate to be extremely close to the interest rate on CBDC. As noted by Bordo and Levin (2017), were the central bank to engage in open market operations between Treasury bills and CBDC, they could ensure that this would be the case in practice.

We follow the same process for the banking sector. Banks can hold assets in the form of CBDC (C), loans (L), or government bonds (B). Again, this is consistent with our balance sheet diagram. As with nonbanks, the banking sector receives both a pecuniary return on its CBDC holdings, R^C, and a nonpecuniary benefit from CBDC's transactional services, ϕ^B. Unlike nonbanks, they receive an additional nonpecuniary benefit, η, from CBDC through its role as a high-quality liquid asset (HQLA). This could be due to a regulatory need to hold HQLA, or a portfolio preference of the bank itself. In a system of mandated reserve requirements, this benefit could be very significant. Government bonds also provide a benefit as HQLAs, ϵ, but provide no transactional services, and default with probability δ.

The last asset that banks can hold on their balance sheet is loans. The pecuniary return is R^L with a default probability of μ. We assume a cost of producing and monitoring each loan, M. For a more developed model which includes monitoring costs of this type, see Goodfriend and McCallum (2007). We further assume that loans offer no transactional services, nor liquidity services. Lastly, banks must finance the asset side of their balance sheet with liabilities, meaning that they must pay R^D on all deposits owed to the nonbank sector.

Taken together, we can write the bank's optimization problem as

$$\begin{aligned} &R^C C + \phi^B C + \eta C + (1-\delta)[R^B + \epsilon]B \\ &\quad + (1-\mu)R^L L - ML - R^D[C+B+L](1-\gamma), \end{aligned} \tag{A.10}$$

the first-order conditions of which give

$$\frac{dU}{dC} = R^C + \phi^B + \eta - (1-\gamma)R^D = 0 \tag{A.11}$$

$$\frac{dU}{dB} = (1-\delta)R^B + (1-\delta)\epsilon - (1-\gamma)R^D = 0 \tag{A.12}$$

$$\frac{dU}{dL} = (1-\mu)R^L - M - (1-\gamma)R^D = 0 \tag{A.13}$$

and which optimizes to give

$$R^B = \frac{R^C + \phi^B + \eta - (1-\delta)\epsilon}{(1-\delta)} \tag{A.14}$$

$$R^L = \frac{R^C + \phi^B + \eta + M}{(1-\mu)}. \tag{A.15}$$

This shows that from the viewpoint of the banking sector, the spread of bond rates over the CBDC rate is a positive function of the transactional benefit of CBDC, the default risk in bonds, and the relative benefits of CBDC as a HQLA when compared with bonds. When combined with the nonbank condition for bonds, this implies that the relative transactional services received by banks compared with nonbanks must equal the additional benefit that banks receive from holding bonds relative to CBDC as HQLA. The loan rate spread is a positive function of the cost of producing a loan and the probability of default.

References

Armelius, H., P. Boel, C. Claussen, and M. Nessen. 2018. "The e-krona and the Macroeconomy." *Economic Review* (Sveriges Riksbank) 2018 (3): 43–62.

Assenmacher, K., and S. Krogstrup. 2018. "Monetary Policy with Negative Interest Rates: Decoupling Cash from Electronic Money." IMF Working Paper No. 18/191.

Bank for International Settlements. 2018. "Central Bank Digital Currencies."

Bank of Canada. 2017. "Digital Currencies and Fintech." Retrieved from https://www.bankofcanada.ca/research/digital-currencies-and-fintech/.

Bank of England. 2015. "Sterling Monetary Framework Annual Report 2014-15." Retrieved from https://www.bankofengland.co.uk/-/media/boe/files/sterling-monetary-framework/annual-report-2014-15.pdf.

———. 2020. "Central Bank Digital Currency: Opportunities, Challenges and Design." Discussion Paper, Bank of England.

Barrdear, J., and M. Kumhof. 2016. "The Macroeconomics of Central Bank Issued Digital Currencies." Staff Working Paper No. 605, Bank of England.

Basten, C., and M. Mariathasan. 2018. "How Banks Respond to Negative Interest Rates: Evidence from the Swiss Exemption Threshold." CESifo Working Paper No. 6901.

Bech, M. L., and R. Garratt. 2017. "Central Bank Cryptocurrencies." *BIS Quarterly Review* (September): 55–70.

Benos, E., R. Garratt, and P. Gurrola-Perez. 2017. "The Economics of Distributed Ledger Technology for Securities Settlement." Working Paper No. 670, Bank of England.

Bjerg, O. 2017. "Designing New Money: The Policy Trilemma of Central Bank Digital Currency." Working Paper, Copenhagen Business School.

———. 2018. "Breaking the Gilt Standard — The Problem of Parity in Kumhof and Noone's Design Principles for Central Bank Digital Currencies." Working Paper, Copenhagen Business School.

Bordo, M., and A. Levin. 2017. "Central Bank Digital Currency and the Future of Monetary Policy." NBER Working Paper No. 23711.

———. 2019. "Digital Cash: Principles and Practical Steps." NBER Working Paper No. 25455.

Broadbent, B. 2016. "Central Banks and Digital Currencies." Speech given at the London School of Economics, March 2.

Butt, N., R. Churm, and M. McMahon. 2015. "Did Quantitative Easing Boost Bank Lending?" Blog Entry, Bank Underground, July 17. Retrieved from https://bankunderground.co.uk/2015/07/17/did-quantitative-easing-boost-bank-lending/.

Carney, M. 2018. "The Future of Money." Speech to the Inaugural Scottish Economics Conference, Edinburgh University, March 2.

Chiu, J., and J. Hill. 2015. "The Rate Elasticity of Retail Deposits in the United Kingdom: A Macroeconomic Investigation." Staff Working Paper No. 540, Bank of England.

Christensen, J., and S. Krogstrup. 2019. "Transmission of Quantitative Easing: The Role of Central Bank Reserves." *Economic Journal* 129 (617): 249–72.

Danezis, G., and S. Meiklejohn. 2016. "Centrally Banked Cryptocurrencies." University College London.

Eichengreen, B. 2019. "From Commodity to Fiat and Now to Crypto: What Does History Tell Us?" NBER Working Paper No. 25426.

Friedman, M. 1960. *A Program for Monetary Stability.* Fordham University Press.

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.

Fung, B. S., and H. Halaburda. 2016. "Central Bank Digital Currencies: A Framework for Assessing Why and How." Staff Discussion Paper No. 2016-22, Bank of Canada.

Goodfriend, M. 2002. "Interest on Reserves and Monetary Policy." *Economic Policy Review* (Federal Reserve Bank of New York) 8 (1): 77–84.

———. 2016. "The Case for Unencumbering Interest Rate Policy at the Zero Bound." In *Designing Resilient Monetary Policy Frameworks for the Future.* Proceedings of the 2016 Economic Policy Symposium sponsored by the Federal Reserve Bank of Kansas City, held in Jackson Hole, Wyoming, August 25–27.

Goodfriend, M., and B. T. McCallum. 2007. "Banking and Interest Rates in Monetary Policy Analysis: A Quantitative Exploration." *Journal of Monetary Economics* 54 (5): 1480–1507.

Hampl, M., and T. Havranek. 2018. "Central Bank Capital as an Instrument of Monetary Policy." Econstor.

He, D., R. Leckow, V. Haksar, T. Mancini-Griffoli, N. Jenkinson, M. Kashima, T. Khiaonarong, C. Rochon, and H. Tourpe. 2017. "Fintech and Financial Services: Initial Considerations." IMF Staff Discussion Note No. 17/05.

Jakab, Z., and M. Kumhof. 2015. "Banks are Not Intermediaries of Loanable Funds — and Why This Matters." Staff Working Paper No. 529, Bank of England.

Keating, T., and M. Machiavelli. 2018. "Interest on Reserves and Arbitrage in Post-Crisis Money Markets." FEDS Notes, March 1.

Kimball, M., and R. Agarwal. 2015. "Breaking Through the Zero Lower Bound." IMF Working Paper No. 15/224.

Kumhof, M., and C. Noone. 2018. "Central Bank Digital Currencies — Design Principles and Balance Sheet Implications." Working Paper No. 725, Bank of England.

Lagarde, C. 2020. "The Future of Money — Innovating While Retaining Trust." Retrieved from https://www.ecb.europa.eu/press/inter/date/2020/html/ecb.in201130~ce64cb35a3.en.html.

McLeay, M., A. Radia, and R. Thomas. 2014. "Money Creation in the Modern Economy." *Quarterly Bulletin* (Bank of England) (Q1): 14–27.

Meaning, J., B. Dyson, J. Barker, and E. Clayton. 2018. "Broadening Narrow Money: Monetary Policy with a Central Bank Digital Currency." Staff Working Paper No. 724, Bank of England.

Nier, E., and H. Kang. 2016. "Monetary and Macroprudential Policies — Exploring Interactions." In *Macroprudential Policy,* Vol. 86, ed. Bank for International Settlements, 27–38. Bank for International Settlements.

People's Bank of China. 2020. "Central Bank: Digital RMB Closed Test Will Not Affect RMB Issuance and Circulation." Retrieved from http://www.gov.cn/xinwen/2020-04/17/content_5503711.htm.

Rogoff, K. 2016. *The Curse of Cash.* Princeton University Press.

Scorer, S. 2017. "Central Bank Digital Currency: DLT, or Not DLT? That is the Question." Blog Entry, Bank Underground, June 5. Retrieved from https://bankunderground.co.uk/2017/06/05/central-bank-digital-currency-dlt-or-not-dlt-that-is-the-question/.

Sveriges Riksbank. 2021. "E-krona." Retrieved from https://www.riksbank.se/en-gb/payments--cash/e-krona.

Turner, A. 2015. "The Case for Monetary Finance — An Essentially Political Issue." Paper presented at the IMF 16th Jacques Polak Annual Research Conference, Washington, DC, November 5–6.

Population Aging and the Macroeconomy*

Noëmie Lisack,[a] Rana Sajedi,[b] and Gregory Thwaites[c]
[a]Banque de France
[b]Bank of England
[c]University of Nottingham

We quantify the impact of demographic change on real interest rates, house prices, and household debt in an overlapping-generations model. Falling birth and death rates across advanced economies can explain much of the observed fall in real interest rates and the rise in house prices and household debt. Since households maintain relatively high wealth levels throughout retirement, these trends will persist as population aging continues. Countries aging relatively slowly, such as the United States, will increasingly accumulate net foreign liabilities. The availability of housing as an alternative store of value attenuates these trends, while raising the retirement age has limited effects.

JEL Codes: E21, E43, E13, J11.

1. Introduction

The population of advanced countries has aged rapidly over the past half-century, with life expectancy and the old-age dependency

*We are grateful to Pavel Brendler, Daisuke Ikeda, anonymous referees, and seminar participants at the Banka Slovenije, Danmarks Nationalbank, the Riksbank, the International Monetary Fund, the Reserve Bank of Australia, the Center for Central Banking Studies, the Bank of England, the EUI Alumni Conference, the Federal Reserve Bank of San Francisco, the Money Macro and Finance Conference, and the DEGIT XXII Conference for their feedback. Remaining errors are ours. This paper reflects the opinions of the authors and not necessarily the views of the Bank of England or its committees, or the Banque de France or the Eurosystem. This paper is an updated version of Lisack, Sajedi, and Thwaites (2017). Author contact: Lisack: Banque de France, 31 rue Croix des Petits-Champs, 75001 Paris, France; e-mail: noemie.lisack@banque-france.fr. Sajedi: Bank of England, Threadneedle Street, London, EC2R 8AH, United Kingdom; e-mail: rana.sajedi@bankofengland.co.uk. Thwaites: University of Nottingham, School of Economics, Sir Clive Granger Building, University Park, Nottingham, NG7 2RD, United Kingdom; e-mail: gregory.thwaites@nottingham.ac.uk.

ratio having already reached unprecedented highs and projected to remain high for several decades. At the same time, long-term real interest rates have been on a downward trend, while house prices and household debt have risen dramatically. This paper quantifies the link between these important trends and examines the wider macroeconomic implications of demographic change.

We build a calibrated neoclassical overlapping-generations (OLG) model to quantify the impact of these factors for advanced economies. We find that the aging of the population in advanced economies can explain around 75 percent of the fall of roughly 210 basis points (bps) in the Holston, Laubach, and Williams (2017) measure of natural real interest rates between 1980 and 2015. Demographic pressures are forecast to reduce real interest rates a further 46 bps by 2050. Furthermore, past falls in interest rates, along with the life-cycle pattern of housing demand, mean that demographics can explain more than four-fifths of the 50 percent rise in real house prices between 1970 and the start of financial crisis. Given that the purchase of housing is predominantly carried out using household credit, these developments also explain the doubling of the household debt-to-GDP ratio over that same period.

Our main findings highlight that the concerns about future rises in real interest rates as baby-boomers retire are largely misplaced, for two main reasons. First, it is the stock of wealth, rather than the flow of saving, that determines the interest rate in neoclassical models, and this stock falls only slowly and partially over the course of retirement. The rise in the share of the population in high-wealth stages of life, starting from around age 50 according to the data, will therefore tend to raise the capital-output ratio even in the absence of any behavioral reaction to higher life expectancy. Therefore, we find that while the retired may dissave, they still hold more wealth, and this higher stock pushes down on the interest rate. This is in contrast to the findings in Carvalho, Ferrero, and Nechio (2016) and the arguments in Goodhart and Pradhan (2017). Second, population aging is predominantly linked to a trend rise in life expectancy, rather than the transient rise in birth rates in the baby boom. The ongoing rise in life expectancy will, other things equal, tend to raise average wealth at any point in life as households anticipate needing to provide for a longer retirement. The aging of the baby-boom generations merely changes the timing of these long-run life-expectancy effects.

While the size and timing of the effects we find are sensitive to model calibration and specification, the mechanism underlying our model is quite general. The rise in life expectancy tends to raise the capital-labor ratio: households save more for longer retirements and spend longer in high-wealth phases of their lives, and this extra wealth finds its way to firms and finances their capital investment. This pushes down on the marginal product of capital, which is a key determinant of the real interest rate.[1] The fall in interest rates may encourage or dissuade further saving, thereby strengthening or weakening the effect of the original demographic shock. This depends on whether households' savings function slopes upward or downward.

We use our model to study the role that housing plays in mediating these effects. If productive capital is the only store of value, all of the burden is placed on capital to meet any increase in desired wealth holdings. So the presence of alternative stores of value can affect the impact of demographics on the interest rate. We show that, in practice, the presence of housing attenuates the fall in interest rates induced by demographic change, although the effects appear to be quantitatively small in our baseline calibration.

Using the heterogeneity by age and birth year within our model, we examine the implications of demographic changes for the level and distribution of welfare across time. Looking at expected welfare across cohorts over time, the main driver is increased longevity, which is conceptually difficult to quantify. Other than longevity, changes in lifetime consumption are the main drivers of lifetime utility, and seem to be detrimental to the baby-boom generations. Lower interest rates tend to increase welfare by allowing agents to better smooth consumption over their life cycles.

Finally, we delve further into the open-economy dimension of our model by considering the demographic transition from the

[1] In fact, in a frictionless neoclassical model such as ours, the marginal product of capital is the only determinant of the real interest rate, while it could be one of several in a model with risk premiums, financial frictions, or price markups (Caballero, Farhi, and Gourinchas 2017; Eggertsson, Mehrotra, and Robbins 2019).

perspective of each country, taking the world interest rate as given. Specifically, while the world interest responds to the aggregate demographic trend, differences in the size and speed of demographic change across countries will lead to capital flows between countries. We show that demographics explain around 20 percent of the dispersion in advanced-economy net foreign asset positions.[2] Given the difficulty of taking such a low-frequency model to time-series data, the explanatory power of the model in the cross-section gives us greater confidence in its predictions about the average level of interest rates in the global economy.

We do not claim that demographic change is the only influence on interest rates and other macroeconomic trends over the long run. In common with the other papers in this literature, our model is very stylized, and in particular does not include an account of the risk premiums and financial frictions that may have caused the return on capital to diverge from government bond yields (see, for example, Caballero, Farhi, and Gourinchas 2017; Marx, Mojon, and Velde 2018). The aim of this study is to isolate the effect of demographic change on savings behavior and the real return on capital and other assets.

Our paper is one of several addressing the impact of demographic change on the real interest rate, house prices, or external payments. Many papers focus only on a single economy: see, for instance, Kiyotaki, Michaelides, and Nikolov (2011), Gagnon, Johannsen, and López-Salido (2021), and Eggertsson, Mehrotra, and Robbins (2019) on the United States or Waldron and Zampolli (2010) on the United Kingdom. Gagnon, Johannsen, and López-Salido (2021) and Eggertsson, Mehrotra, and Robbins (2019) model in detail the household's life cycle, but relative to this paper they focus on the role of demographic change on the decline in the natural interest rate in the United States, while Kiyotaki, Michaelides, and Nikolov (2011), conversely, focus on the effect of interest rates and demographics as drivers of land and housing markets. By considering all advanced economies together, instead, our paper relates to Carvalho,

[2]Net foreign asset positions in the data are generally smaller in magnitude than the values predicted by our model, since the model ignores the presence of frictions in international capital flows.

Ferrero, and Nechio (2016) and Marx, Mojon, and Velde (2018) and contributes to that literature by using a fully fledged life-cycle model, examining the effects of demographics on house prices and household credit, as well as accounting more precisely for the high levels of wealth observed among older households, and highlighting the effect of this stock of wealth on the interest rate. Last, our open-economy exercise builds on Domeij and Flodén (2006), Krueger and Ludwig (2007), and Backus, Cooley, and Henriksen (2014), who study international capital flows via current account movements. It is also related to the more recent work of Barany, Coeurdacier, and Gribaud (2018), who take into account country-specific borrowing constraints and social security. Here as well, our paper contributes to the literature by including housing as an alternative store of value, and by highlighting the importance of wealth stocks—and not only flows—for evaluating the impact of demographic change on the macroeconomy.

The remainder of this paper is structured as follows. Section 2 sets out the key demographic trends over the past few decades. Section 3 describes the model and its calibration. Section 4 shows the results of model simulations in which we incorporate the demographic trends, and considers some robustness exercises. Section 5 concludes.

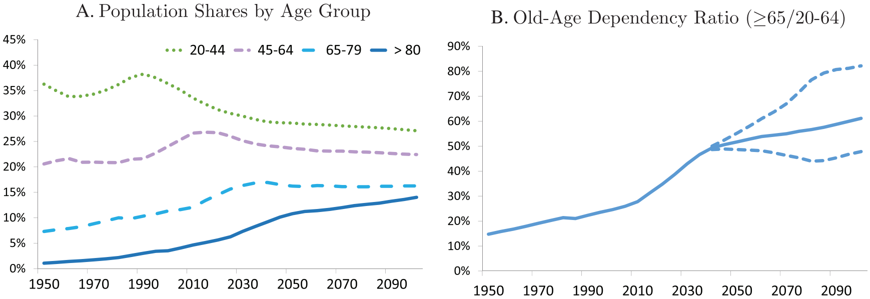
2. Demographic Trends

To document the key demographic trends that motivate this paper, we use data from the UN Population Statistics, which runs from 1950 to 2015 and includes projections up to 2100. Our focus is on an aggregate of advanced economies.[3]

The shares of different age groups in the total population over time, shown in figure 1A, present two main patterns. Firstly, we see a clear rise in the share of older generations, for example, with the over-80s going from 1 percent of the population in 1950 to around 5 percent in 2015 and reaching a projected 14 percent by the end of

[3]In particular, we use Western Europe (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom), North America (Canada and the United States), Japan, Australia, and New Zealand.

Figure 1. Aging Population



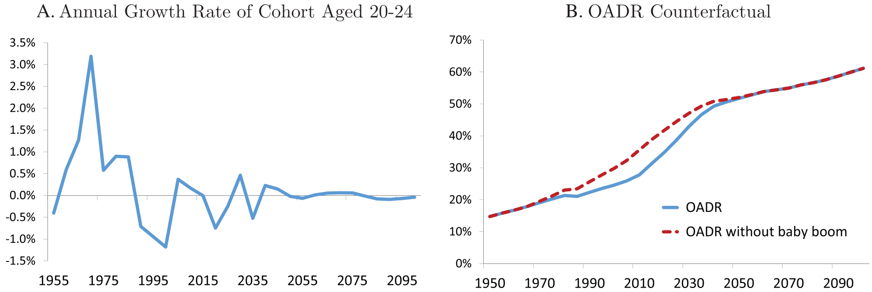
Source: UN Population Statistics (projections based on medium-fertility scenario; dashed lines in panel B show high- and low-fertility scenarios).

the century. Secondly, the effect of the baby boom shows as a "bulge" moving through the population, entering the 20–44 age group from the 1970s and slowly disappearing by around 2040.

This evolution is summarized by the old-age dependency ratio (henceforth OADR), shown in figure 1B, defined as the ratio of over-65s to 20- to 64-year-olds. Again the clear upward trend shows the rise in the share of older generations relative to the working population. The dashed lines, which show the alternative fertility scenarios in the UN projections, give an indication of the degree of certainty around the projections: even in the high-fertility scenario, the OADR increases substantially from around 15 percent in 1950 to almost 50 percent in 2100. In the medium-fertility scenario the final number is over 60 percent.

Having documented these trends, we now examine their causes. Figure 2A shows the growth rate of consecutive 20- to 24-year-old cohorts over time. We can see that the period 1970–85 saw elevated growth rates, reaching over 3 percent per annum, corresponding to the baby-boom generations born between 1945 and 1960. Growth rates have since fallen, and have even been significantly negative for several periods. Both of these affect the age structure of the population. In particular, as the large baby-boom cohorts grow old, the age distribution becomes skewed toward the older age groups. This is amplified by the smaller size of the new younger generations entering the population.

Figure 2. The Baby Boom



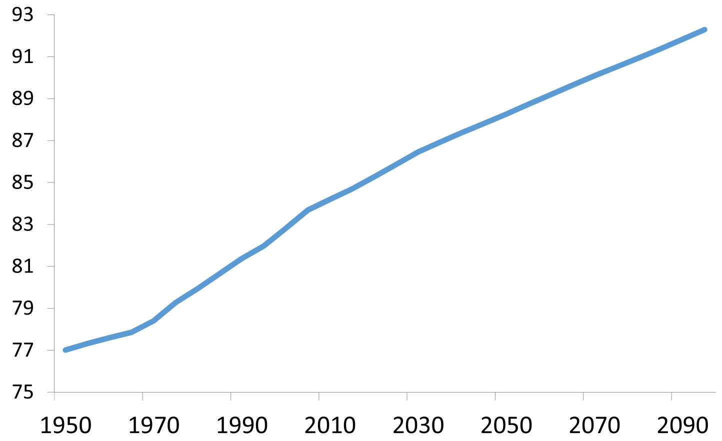
Source: UN Population Statistics (projections based on medium-fertility scenario).

To further illustrate the baby-boom effect on the aggregate demographic trends, figure 2B shows the counterfactual OADR when we assume that cohort growth in 1970–85 was zero, hence removing the effect of the baby boom. We can see that there is a non-negligible effect from these cohorts. When they are young, and on the denominator of the OADR, they lower this ratio relative to the counterfactual. As they get older and begin to move to the numerator of the ratio, they account for a steeper rise in the OADR. Nonetheless, once these cohorts have faded out of the population, the counterfactual OADR reaches the same high levels as the baseline projections. Hence the baby boom does not account for the long-run trend in the OADR.

The key determinant of the rise in the OADR is increasing longevity. Figure 3 shows life expectancy conditional on living to age 60.[4] While a 60-year-old in 1950 would not expect to live past the age of 77, by 2015 a 60-year-old can expect to live until close to 85. By the end of the century this number rises past 90. As people face lower mortality rates later in life, and their life expectancy rises, older age groups account for a growing proportion of the total population.

[4]This measure is taken directly from the UN Population Statistics, and is defined as "the average … years of life expected by a hypothetical cohort of individuals alive at age 60 who would be subject during the remaining of their lives to the mortality rates of a given period."

Figure 3. Life Expectancy at 60



Source: UN Population Statistics (projections based on medium-fertility scenario).

As the data make clear, aging population in advanced economies has led to an unprecedented shift in the age structure of the population, and these effects will almost certainly persist for decades to come. The rest of this paper will employ an overlapping-generations model to uncover the macroeconomic effects of these important trends.

3. Quantitative Model

3.1 The Model

We consider a general equilibrium setup with overlapping-generation households and a representative firm producing in a perfectly competitive environment.[5] We describe these two agents in turn, and then describe the aggregation and market clearing conditions.

[5]The underlying model is the same as that of Sajedi and Thwaites (2016), allowing for demographic change.

3.1.1 Household

Agents are born at age 1 and can live up to T periods. We denote by $x_{\tau,t}$ the value of a variable x, for a household born in period t, when they are aged τ.

Agents work from their first period of life until they reach retirement, at a fixed age T^r. They face a probability of death at (after) each age τ, denoted $(1-\psi_{\tau,t}) > 0$, and die with certainty at the maximum age, T, hence $\psi_{T,t} = 0 \,\forall t$. This can be translated into the probability of surviving until each age, $\tilde{\psi}_{\tau,t} = \Pi_{j=1}^{\tau-1}\psi_{j,t}$, with $\tilde{\psi}_{1,t} = 1$.

Throughout their life, agents supply labor, l, inelastically, and gain utility from a consumption good, c, and a housing good, h, which is traded at relative price p^h. Hence the T-period optimization problem faced by a representative household born in period t can be written as

$$\max_{\{c_{\tau,t},\, a_{\tau,t},\, h_{\tau,t}\}_{\tau=1}^{T}} \quad \sum_{\tau=1}^{T} \beta_\tau \tilde{\psi}_{\tau,t} \left(\ln c_{\tau,t} + \theta_\tau \ln h_{\tau,t}\right) + \beta_T \tilde{\psi}_{T,t} \phi \ln a_{T,t}$$

subject to

$$\begin{aligned} & c_{\tau,t} + a_{\tau,t} + p^h_{t+\tau-1}(h_{\tau,t} - h_{\tau-1,t}) \\ & \quad \leq w_{t+\tau-1}\epsilon_\tau l_{\tau,t} + (1 + r_{t+\tau-1})a_{\tau-1,t} + \pi_{\tau,t} \qquad \text{for } \tau = 1, \ldots, T, \end{aligned}$$

where ϵ is the age-specific productivity level, w is the wage per efficiency units of labor, and a is a safe asset with return r.[6] We assume that $l_{\tau,t} = \epsilon_\tau = 0$ for $\tau \geq T^r$.

Agents are born without any assets, that is, $a_{0,t} = 0$, but we allow the possibility of bequests, setting $\phi > 0$ so that $a_{T,t} > 0$. These bequests are distributed among subsequent generations as part of $\pi_{\tau,t}$, which captures all nonlabor income, taken as exogenous by the households.

There are a fixed number of periods when the household is able to "move house," i.e., reoptimize their housing wealth, and hence

[6]Note that $t + \tau - 1$ is the period in which the generation born at time t is aged τ.

outside of these "move dates" the household has an additional constraint $h_{\tau,t} = h_{\tau-1,t}$.[7] We assume that agents are born without any housing wealth and do not leave any housing wealth when they die, hence $h_{0,t} = h_{T,t} = 0$, which necessitates $\theta_T = 0$.

Denoting by $\lambda_{\tau,t}$ the Lagrange multiplier on the budget constraint at age τ, this problem gives rise to the following first-order conditions:

$$\lambda_{\tau,t} = \beta_\tau \tilde{\psi}_{\tau,t} c_{\tau,t}^{-1} \qquad \forall \tau = 1, \ldots, T$$

$$\lambda_{\tau,t} = (1 + r_{t+\tau})\lambda_{\tau+1,t} \qquad \forall \tau = 1, \ldots, T-1$$

$$\lambda_{T,t} = \beta_T \tilde{\psi}_{T,t} \phi a_{T,t}^{-1}$$

$$\sum_{j=\tau}^{\tau'-1} \beta_j \tilde{\psi}_{j,t} \theta_j h_{\tau,t}^{-1} = p_{t+\tau-1}^h \lambda_{\tau,t} - p_{t+\tau'-1}^h \lambda_{\tau',t} \qquad \forall \tau \in \text{"move dates,"}$$

where τ' in the last equation denotes the next move date after τ.

3.1.2 Firm

The firm's problem is to choose the aggregate factors of production, K_t and L_t, to maximize profit, taking as given the rental rate of capital, r_t^k, the wage per efficiency units of labor, w_t, and the production function, $Y = F(K, L)$. Note that L_t denotes the aggregate efficiency units of labor supplied by households. This problem can be written as

$$\max_{L_t,\, K_t} \quad F(K_t, L_t) - w_t L_t - r_t^k K_t.$$

Taking the constant elasticity of substitution (CES) production function $F(K, L) = \left[(1-\alpha)L^{\frac{\sigma-1}{\sigma}} + \alpha K^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$, we have the following first-order conditions[8]:

---

[7] This is a simple way to capture the fixed costs associated with buying and selling housing, which makes housing a less liquid asset.

[8] Although changes in total factor productivity (TFP) growth over time may partly explain changes in real interest rates, we abstract from this in order to focus on the role of demographics. Instead, below, we will look at how demographic changes affect labor productivity due to the age-specific productivity levels.

$$w_t = (1-\alpha)\left(\frac{Y_t}{L_t}\right)^{\frac{1}{\sigma}}$$

$$r_t^k = \alpha\left(\frac{Y_t}{K_t}\right)^{\frac{1}{\sigma}}.$$

Capital is financed from the households savings and depreciates at rate δ every period. Before paying for the capital rental rate, the firm is left with $(1-\delta+r_t^k)K_t$ at the end of each period t and the households receive an interest rate r_t on their savings, hence the zero-profit condition of the firm implies

$$r_t^k = r_t + \delta.$$

3.1.3 Aggregation

We denote the gross growth rate of the generation born at time t relative to the generation born at time $(t-1)$ with g_t. Normalizing the size of the generation born at time 0 to 1, this means the size of the generation born at time t can be written as

$$s_t = g_t s_{t-1} = \Pi_{i=1}^{t} g_i.$$

At each age, the size of the cohort reduces, with survival probability $\tilde{\psi}_{\tau,t} \leq 1$. Hence the total population in period t is given by

$$S_t = \sum_{\tau=1}^{T} \tilde{\psi}_{\tau,t-\tau+1} s_{t-\tau+1} = \sum_{\tau=1}^{T} \tilde{\psi}_{\tau,t-\tau+1} \Pi_{i=1}^{t-\tau+1} g_i.$$

Let $\mathbf{x}_t$ denote the $(T \times 1)$ vector of a variable x, for one representative household of each generation alive at time t; in other words, $\mathbf{x}_t = \{x_{\tau,t-\tau+1}\}_{\tau=1}^{T}$. Let ρ_t denote the $(T \times 1)$ vector of population sizes at time t, that is, $\rho_t = \{\tilde{\psi}_{\tau,t-\tau+1} s_{t-\tau+1}\}_{\tau=1}^{T}$. The aggregate value of variable x at time t is denoted by $X_t = \rho_t' \mathbf{x}_t$. We denote by $\tilde{X}_t$ the value of X_t per aggregate capita, that is, X_t/S_t. We can write this as $\tilde{X}_t = \tilde{\rho}_t' \mathbf{x}_t$ where $\tilde{\rho}_t = \rho_t/S_t$ denotes the vector of *relative* population sizes.

3.1.4 Market Clearing

Capital/Savings Market. The value of the capital stock must equal the aggregate savings of the previous period,

$$A_{t-1} = K_t.$$

As introduced above, we denote per capita capital stock as $\tilde{K}_t = K_t/S_t$. For consistency, this implies that per capita savings are defined relative to next period's population, that is, $\tilde{A}_t = A_t/S_{t+1} = \frac{S_t}{S_{t+1}}\tilde{\rho}_t'\mathbf{a}_t$.

Labor Market. Aggregate labor supply must equal labor demand. Let $\epsilon\mathbf{l}_t = \{\epsilon_\tau\, l_{\tau,t-\tau+1}\}_{\tau=1}^T$ denote the vector of efficiency units of labor supplied by each generation at time t. Then

$$\rho_t'\epsilon\mathbf{l}_t = L_t \quad \Rightarrow \quad \tilde{\rho}_t'\epsilon\mathbf{l}_t = \tilde{L}_t.$$

Housing Market. As with the household savings, for consistency we define per capita housing relative to next period's population, that is, $\tilde{H}_t = H_t/S_{t+1} = \frac{S_t}{S_{t+1}}\tilde{\rho}_t'\mathbf{h}_t$. Housing is effectively residential land, in that its supply is inelastic, hence we assume that the housing stock per capita is fixed at some level, $\tilde{H}$.[9] Market clearing then simply requires

$$\tilde{H}_t = \tilde{H} \quad \forall t.$$

This implies that the aggregate housing stock, H_t, grows with the population, meaning that the economy is endowed with an additional $(\frac{S_{t+1}}{S_t} - 1)\tilde{H}$ units of housing each period.[10] This endowment is distributed across households through nonlabor income, along with the bequests, as detailed below.

Bequests and Nonlabor Income. At each time t, the nonhousing assets and the housing wealth of the generations that died

[9]This is in line with Knoll, Schularick, and Steger (2017), who find that the bulk of the increase in house prices is attributable to the increase in the value of residential land.

[10]The alternative would be to allow this additional housing to be produced, with a technology which transforms the consumption good into housing. This does not materially affect the results, as discussed in section 4.3 and in online appendix C.1, available at http://www.ijcb.org.

in the previous period, as well as the additional housing endowment, added in each period to maintain a stable level of housing per capita, are distributed to living households through bequests (see online appendix A.1, at http://www.ijcb.org, for more details). This nonlabor income is evenly distributed among households above a given age, T^b, while younger households are not entitled to any nonlabor income.[11]

Goods Market. Aggregating the budget constraints of all households alive at a given time t, and substituting the equilibrium conditions described above, gives us the familiar resource constraint

$$\tilde{Y}_t = \tilde{C}_t + \tilde{I}_t,$$

where $\tilde{I}_t$ is the net increase in aggregate savings, given by

$$I_t = A_t - (1-\delta)A_{t-1} \quad \Rightarrow \quad \tilde{I}_t = \frac{S_{t+1}}{S_t}\tilde{A}_t - (1-\delta)\tilde{A}_{t-1}.$$

Hence the resource constraint above simply implies that all goods produced at time t are either consumed or saved as capital.

3.2 Calibration

Each period in the model represents five years. We assume that working life begins at age 20 and no agents live beyond age 90, setting $T = 14$.

The focus of our calibration will be (i) to match life-cycle profiles of labor productivity, housing wealth, and net worth, and (ii) to match aggregate housing wealth-to-GDP, debt-to-GDP, and real interest rates. All of these moments will vary over time in the dynamic transition path due to the demographic trends. Hence we must target these moments at particular points in time. Given the data availability, we target average life-cycle patterns for the years 1990–2010. For the aggregate moments, we target their average values over the 1970s, in order to allow the model to determine the

[11] This assumption aims to reflect the fact that older households are more likely to see their family members die and to inherit their assets and housing wealth. Furthermore, a flat bequest distribution across households above this age ensures that bequests do not create strong distortions on the household consumption and saving choices.

transition over the past few decades. Full details of the calibration procedure are provided in online appendix A.2.

3.2.1 Demographic Transition

Population growth, g_t, and the survival probabilities, $\psi_{\tau,t}$, are the exogenous demographic processes that drive fluctuations in our model. Using the data described and shown in section 2, we set these two series so as to match the evolution of the age structure of the economy from the 1950s, and projected until 2100.

Specifically, we set g_t as the relative size of consecutive 20- to 24-year-old cohorts over time. We then set $\psi_{\tau,t}$ to match the observed evolution of each cohort throughout their life, meaning that the rate of decline in the size of a given cohort from one period to the next is taken to be the death rate.[12]

3.2.2 Data

Life-Cycle Profiles. Given limited cross-country data availability, we will assume that U.S. households are representative of all advanced-economy households in terms of the life-cycle profiles of labor productivity, housing wealth, and net worth. Hence, we can use the Survey of Consumer Finance (SCF) to match life-cycle profiles for productivity, ϵ, net worth, a, and housing wealth, h.

Specifically, we calibrate productivity to match *"Wage Income"* data from the SCF, which corresponds to total labor income, irrespective of hours worked. Hence, since hours worked are inelastic in the model, we are effectively subsuming all life-cycle hours and wage decisions into the productivity profile. To calibrate housing wealth over the life cycle, we take the sum of *"Primary Residence"* and

[12]The existence of immigration means that cohort sizes can go up as well as down over time, particularly for younger age groups. To remove this possibility, we smooth the death rate before retirement to match the overall decline of a given cohort between the ages of 20 and 64. If a cohort size is higher at the age of 64 than at the age of 20, which is the case for more recent years, we assume a zero probability of death before retirement.

"Other Residential Real Estate" in the SCF. We then subtract this from the total *"Net Worth"* to obtain nonhousing assets a.[13]

Aggregate Variables. We take three aggregate variables as targets: the real interest rate, housing wealth-to-GDP, and debt-to-GDP. In order to allow the model to determine the evolution of these variables over the last few decades, we target their average value in the 1970s.

For the real interest rate we use the data from Holston, Laubach, and Williams (2017), and take the average world interest rate between 1970 and 1980. This gives us a target of 3.42 percent.

The data from Piketty and Zucman (2014) measure housing assets, including land, and give us the aggregate housing wealth-to-GDP target. We take an average over the 1970s for all available countries, namely Australia, Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States. We obtain a target ratio of 147 percent.

Finally, for debt-to-GDP we use the Bank for International Settlements (BIS) Total Credit data, focusing on total credit to households as a percentage of GDP. Again we use the average over the 1970s for the countries available—in this case, Canada, Germany, Italy, Japan, the United Kingdom, and the United States. The final target is 40 percent.

3.2.3 Other Parameters

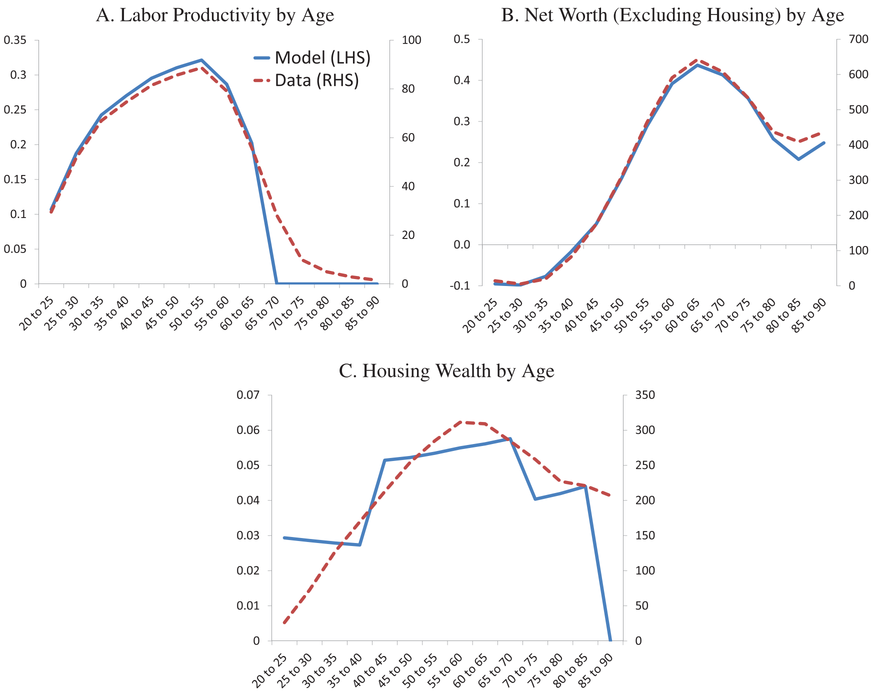
We set the parameters of the CES production function $\sigma = 0.7$ and $\alpha = 1/3$, and the annualized depreciation rate $\delta = 6\%$.

We set hours worked at 0.3 throughout working life, hence $l_\tau = 0.3$ for $\tau = 1, \ldots, T^r - 1$, and $l_\tau = 0$ for $\tau \geq T^r$. Hence aggregate labor supply is $L = 0.3$, the value commonly used in the literature. Households are assumed to retire at age 65, corresponding to $T^r = 10$. They start receiving bequests at age $T^b = 7$, i.e., age 50.

In matching the life-cycle profile of housing wealth, we set the "move dates" in the household's problem to $\tau = 1, 5, 11$, corresponding to ages 20, 40, and 70. The final period of life will also be

[13] See online appendix A.3 for details. Note that by matching labor income and assets, consumption over the life cycle will be determined residually by the household budget constraint, and so it is not being matched to data.

Figure 4. Calibration of Life-Cycle Profiles



Notes: Model lines are relative to the expected lifetime labor income of the 1950 cohort. Data lines are thousands of U.S. dollars.

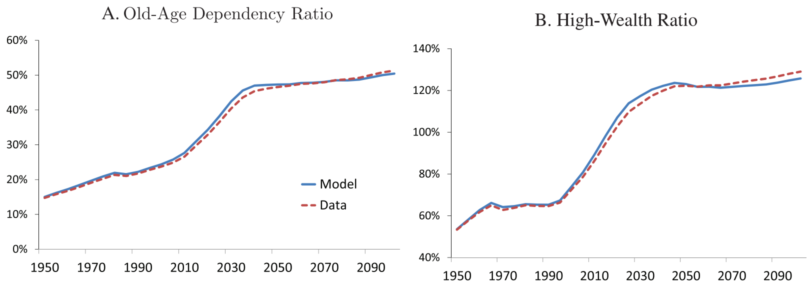
a "move date," as we have assumed that agents do not leave housing bequests.

3.2.4 Calibration Outcomes

Figure 4 shows the life-cycle profiles for productivity, housing, and net worth from the data and the model. For the model, given that these profiles change over time in the transition, we take the equivalent of the estimates from the data, namely the average of the cross-sectional age profile of each variable over 1990–2015. The aggregate moments are matched exactly by construction.

To show how well we fit the demographic trends with the growth rate and death probabilities as exogenous series, figure 5A shows the

Figure 5. Demographics in the Model vs. Data



Source: UN Population Statistics and own calculations.

OADR of the model against the data. Figure 5B plots a slightly different ratio, which we call the high-wealth ratio (HWR). To define this ratio, we use the empirical life-cycle profile of assets to define the "high-wealth" phase of an agent's life. As can be seen from figure 4B, agents have accumulated a large amount of wealth by around the age of 50–55, and maintain that level of wealth until the end of their life at the age of 90. Hence we define the HWR as the ratio of those over 50 to those aged 20–49. Looking at figure 5, we see that, despite the simplifications that we make, both the OADR and the HWR in the model are very close to that in the data.[14]

4. Results

We now present the results of the model simulations. Given the exogenous demographic changes described above, we solve for the general equilibrium transition path of the economy, assuming perfect foresight. We show the transition of the main macroeconomic variables of interest—namely the real interest rate, savings, and debt—and decompose these results in terms of the changes in the age distribution of the population and changes in the savings behavior of households as they expect to live longer. We then explore other

[14]Note that the data line here ignores the population over 90, in line with the model, and so the OADR does not rise as much as in figure 1.

macroeconomic implications of demographic trends: on the housing market, labor productivity, welfare and distribution, and in an open-economy context. To facilitate the comparison with the model outcomes, all data points shown in the figures are five-year averages. Finally, we carry out some robustness exercises to further investigate our results.

4.1 Baseline Results

4.1.1 The Interest Rate, Savings, and Debt

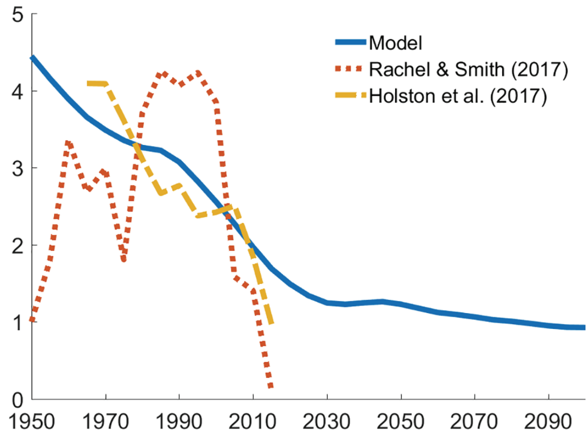
First, we turn to the main outcome of our model, namely the real interest rate, compared with its empirical counterparts. Given that we are using a real model, the real interest rate here should be interpreted as the natural real interest rate: the interest rate that prevails in the absence of nominal rigidities or, equivalently, the interest rate that is consistent with inflation at its target and the output gap closed over the medium to long run. This object is not directly observable, but in figure 6 we show the model outcome against two empirical estimates that are available for advanced economies. Namely, the red dotted line shows the average advanced-economy long-run real interest rate, based on 10-year sovereign yields, taken from Rachel and Smith (2017), and the dashed-dotted yellow line shows the model-based estimate of the natural interest rate from Holston, Laubach, and Williams (2017).[15] These measures of real interest rates are clearly more volatile than the long-term natural interest rate that is captured by the demographic trends in our model.

In the model, the annual interest rate decreases by 157 bps between 1980 and 2015.[16] Compared with the empirical counterparts, between 1980 and 2015, demographics are able to explain 75 percent of the roughly 210 bps drop estimated by Holston, Laubach, and Williams (2017) and around 45 percent of the fall in the Rachel

[15]For figures in color, see the online version of the paper at http://www.ijcb.org.

[16]Since we have calibrated the level of the real interest rate in the 1970s, we focus here on the changes since 1980.

Figure 6. The Real Interest Rate (%)

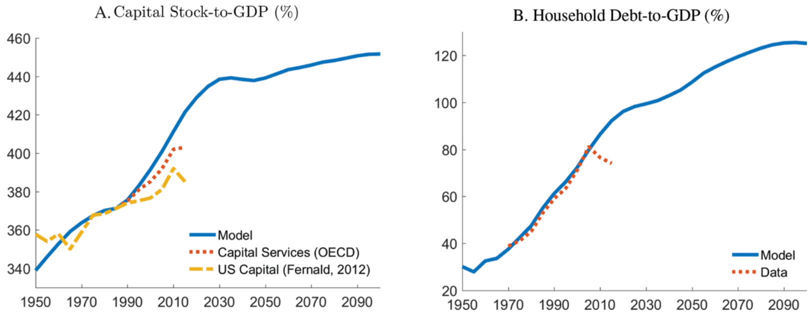


Notes: Holston, Laubach, and Williams (2017) measures the natural interest rate, and is averaged across the United States, United Kingdom, euro area, and Canada. Rachel and Smith (2017) measures the world real interest rate.

and Smith (2017) measure. These empirical measures have fallen by more than is predicted by our model through demographic changes, leaving room for other more transitory explanations of the current low level of interest rates. Nonetheless, it is important to note that the demographic changes themselves do not reverse, and leave the economy with a permanently lower natural interest rate. In the transition path, it is still possible to see the transitory impact of the baby boom, slowing down the interest rate decrease in the 1990s and accelerating it between 2010 and 2040. However, in the long run, the main driver behind the transition path is the increase in life expectancy, as mentioned in section 2, and this trend is projected to persist. The real interest rate in our model is forecast to decrease by a further 46 bps by 2050 and 76 bps by 2100.

The converse of the fall in the real interest rate is a rise in the capital stock, shown in figure 7A against two empirical measures of capital intensity: an index of capital services-to-GDP for 19 advanced countries, in the red dotted line, and an index of the ratio of the capital stock to value added in the U.S. business sector from the Fernald (2012) data, in the yellow dashed-dotted lines.

Figure 7. Capital and Debt

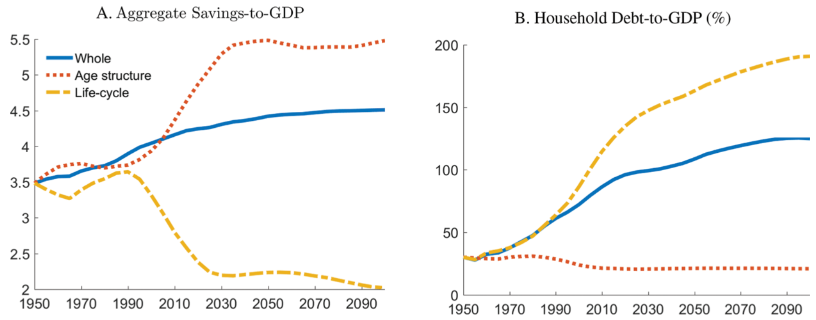


Sources: Capital services are measured by OECD data for Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Korea, Netherlands, New Zealand, Portugal, Spain, Sweden, United Kingdom, and the United States. U.S. capital is from Fernald (2012). Household debt is from BIS databases; the household debt-to-GDP is the ratio of household debt to nominal GDP.

The demographic trends alone slightly overestimate the rise in capital, particularly for the United States, as we abstract from other factors that drive investment. Nonetheless, both measures of capital intensity show a rise, particularly in the period of interest since the 1980s. This gives credence to the notion that part of the fall in the real interest rate is the standard neoclassical effect of a rise in the capital intensity of the economy.

The key mechanisms triggered by the demographic transition are the following. Firstly, households perfectly anticipate that they will live longer and spend more time in retirement. They are therefore willing, all things equal, to transfer more of their income during working life to the future, in order to smooth their consumption. Secondly, the slower population growth and increased longevity imply that older households make up a larger share of the total population alive at each period. Were households not anticipating their longer lifetime and adjusting their consumption and saving choices, this change in population weights would still change aggregate savings and aggregate consumption outcomes. In partial equilibrium, these two mechanisms would both increase the level of aggregate savings-to-GDP over time. In general equilibrium, to keep the capital market balanced given this higher capital supply, the interest rate decreases.

Figure 8. Decomposing the Drivers



Note: Age structure is changing only the population age structure and life cycle is changing only the household's optimal behavior.

This lower interest rate has an offsetting effect, as it discourages savings and even encourages more borrowing by the young, raising net household debt-to-GDP and pushing down on aggregate savings-to-GDP. As shown in figure 7B, household debt-to-GDP rises in the model in line with the data.

We can see the effects of these two mechanisms by decomposing the changes in aggregate savings into the two distinct drivers: changes in the age composition of the population and changes in the life-cycle savings decisions of each household, given the new general equilibrium prices. The impact of these two distinct drivers on aggregate savings and household debt is shown in figure 8.[17] The marginal impact of changes in the population age structure (shown in the red dotted line) on aggregate savings per capita tends to be larger than the baseline: indeed, it only takes into account the smaller (resp. larger) share of younger (resp. older) households in the total population. Since older households hold more assets, increasing their share in the economy drives up the level of aggregate savings per capita. Similarly, only younger households are indebted, so that the aggregate household debt-to-GDP decreases with the decreasing share of young households in the economy.

[17]See online appendix B.1 for the details of the construction of this decomposition and a second exercise distinguishing between partial and general equilibrium effects on household decisions.

Conversely, taking only changes in optimization over the life cycle into account, shown in the dashed-dotted yellow line, the aggregate savings per capita actually decrease massively from 1990 onwards: since the interest rate is lower, and despite their increased life expectancy, households shift their portfolio toward consuming more, holding more housing wealth and more debt when young. Without the offsetting effect of the falling share of the increasingly indebted young in the population, this leads to a fall in aggregate savings.

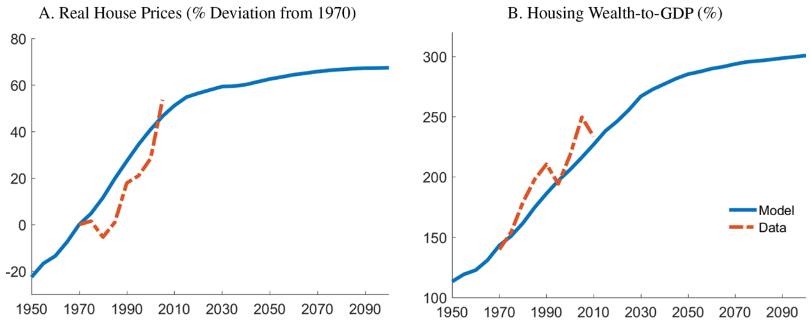
We conduct a similar decomposition exercise for household debt-to-GDP, as shown in figure 8B. We reach similar conclusions to the aggregate savings case: the changes in the population structure tend to stabilize or decrease the household debt-to-GDP, while the household reoptimization of their consumption, savings, and borrowing decisions given their increased life expectancy and the lower interest rate implies a higher debt-to-GDP ratio.

4.1.2 Housing

One important feature of our model compared with the literature is the presence of housing. Households directly derive utility from housing, but housing also serves a second purpose, as households can use it as an additional way of transferring wealth over time, in that it is durable and can be sold to fund consumption and bequests. In our framework, households have perfect foresight, which allows them to anticipate the evolution of the housing price over their lifetime, and therefore anticipate the return on housing as a store of wealth over their life cycle. As the interest rate falls, so does the user cost of housing, which is the opportunity cost of investing into an additional unit of housing instead of the financial asset, and so demand for housing rises. With the supply of housing per capita held fixed in our model, housing prices are pushed up, and, as a consequence, the housing wealth-to-GDP ratio increases, as shown in figure 9. In fact, we are able to explain 85 percent of the observed increase in real house prices.[18] To be able to afford the more expensive housing assets, young households have to borrow more, and so the rising house price also contributes to the rising debt-to-GDP ratio.

[18] This can be seen as an upper bound on this effect, since any increase in housing supply per capita would mitigate the rise in house prices.

Figure 9. Housing in the Baseline Simulations



Sources: Housing wealth comes from Piketty and Zucman (2014). Real house prices are from the BIS databases and are the ratio of nominal house prices to the consumer price index.

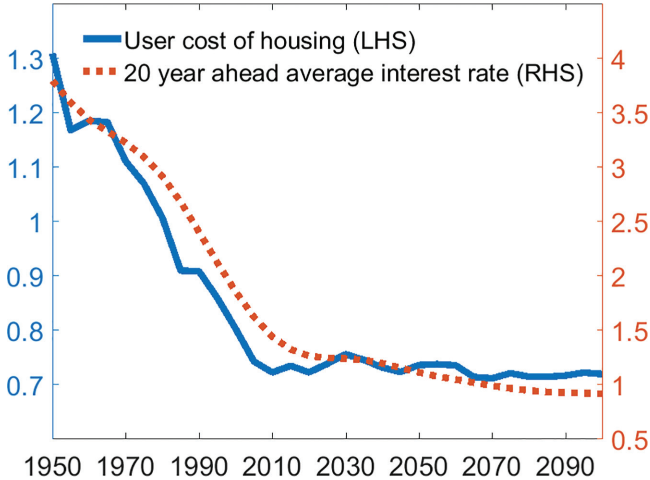
Housing accordingly provides an alternative vehicle for the transfer of resources over the life cycle, and will raise interest rates.

Still, the evolution of house prices does not follow the real interest rate one for one. In our model, the households are only allowed to "move" (change their housing wealth) at specific ages. This has two implications. Firstly, it means that the house price is sensitive to changes in the relative size of different cohorts as they move from buying to selling housing. This means that the baby boom plays an important role in dynamics of house prices. In 2015, the oldest households from the baby-boom generations (born between 1945 and 1949) reach age 70, so that the share of the group aged 40 to 69 in the total population starts decreasing, while the share of the group aged 70 to 84 picks up. Concretely, the share of households with the highest housing demand, the 40- to 69-year-olds, decreases, while the share of households with a lower housing demand, aged 70 and over, increases. Consequently, even as the demographic transition continues, the rise in housing demand slows down around this time, and hence house prices flatten out.[19]

The second implication of the discrete move dates is that, as a savings instrument, housing has a longer maturity than capital, meaning that house prices are more forward looking: households

[19] See online appendix C.1 for more details.

Figure 10. User Cost of Housing and Forward-Looking Interest Rate



Note: The user cost of housing is the product of the interest rate and the house price, less the capital gain on the house.

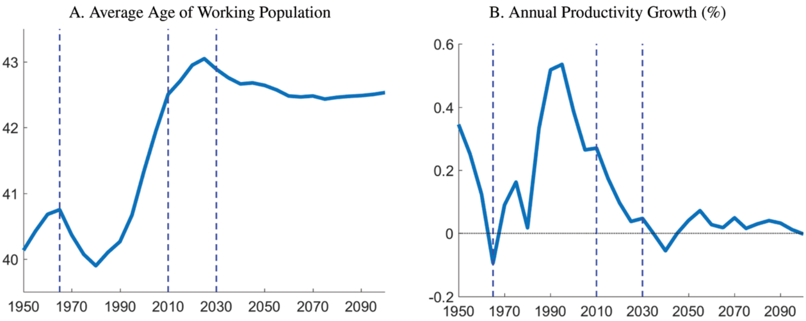
buying housing know that they will have to wait between 15 and 30 years before being able to sell it. The user cost of housing, therefore, reflects expected future changes in prices (both the interest rate and the housing price). This is shown in figure 10, which plots the four-period-ahead (20-year-ahead) average real interest rate against the per-period user cost of housing. These variables follow the transition of house prices more closely than the real interest rate.

4.1.3 Productivity

While we abstracted from trend productivity growth as a driver of macroeconomic trends, our model can partially explain the observed movements in labor productivity through the effect of demographic changes. As shown in figure 4A, the productivity of young and old workers is lower, and productivity reaches its peak around age 50. Hence, a change in the age distribution of the working population implies a different level of the aggregate productivity.

The evolution of the average age of the working population in our model and the resulting productivity growth rate are shown in

Figure 11. Implications of Demographic Trends for Productivity



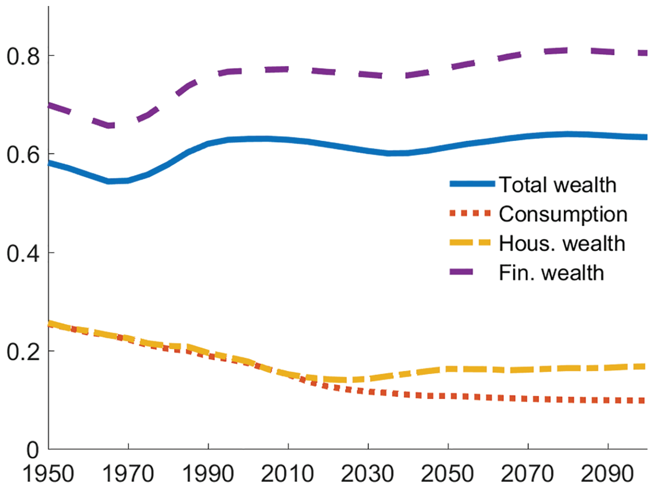
Note: Vertical dashed lines show when the oldest baby-boom generation enters the workforce, retire, and reach age 90.

figure 11. We can clearly see the impact of the baby-boom generations on the figures. From 1970 onwards, the young baby-boomers start working, bringing down the average age of the working population and hence productivity growth. Until 2000, the baby-boomers age and gain in work experience, increasing the labor force's average age and productivity. From 1990, the baby-boomer generations reach ages 50 and above, their productivity decreases, and hence the productivity growth slows down, while the average age of the working population keeps increasing. Finally, from 2015 onwards, the baby-boom generations start retiring progressively. The average age of the working population decreases slowly, and productivity growth slows down further. While demographic changes are not the only explanation for the recent slowdown in productivity growth, our model shows that the aging workforce may have played a role in this evolution.

4.1.4 Distributional Impact of Demographic Change

Since our model includes heterogeneity in terms of birth year and age, we can measure inequality along two dimensions: across households within period, and across cohorts defined by their birth year. It is important to emphasize, however, that our model is not designed

Figure 12. Gini Coefficients



to analyze inequality from either a positive or a normative angle—there is no within-cohort heterogeneity, and there are conceptual difficulties in comparing welfare over changing lifespans.[20] Our results accordingly should not be read as a full intergenerational welfare analysis. Nonetheless, it is interesting to see what insights we can get from the model about distributional impact of demographic changes.

Within-Period Inequality. Figure 12 shows the evolution of Gini coefficients for consumption and financial, housing, and total wealth implied by our model from 1950 onwards. As above, there are two main drivers of this evolution: the direct effect of the changing age structure and the equilibrium effect of the changing life-cycle profiles of each variable. For consumption, these two components work in the same direction. As agents become older, the age structure is more concentrated, lowering the dispersion of consumption and hence lowering the Gini coefficient. At the same time, the lower interest rate allows households to borrow and consume more when young, while older households tend to consume less as they expect

[20]Furthermore, we use model-implied rather than actual relative prices in order to isolate the impact of demographic change, and ignore, as elsewhere in this paper, the impact of other drivers in the data.

to live for longer. Again this lowers the dispersion of consumption across agents and lowers the Gini coefficient.

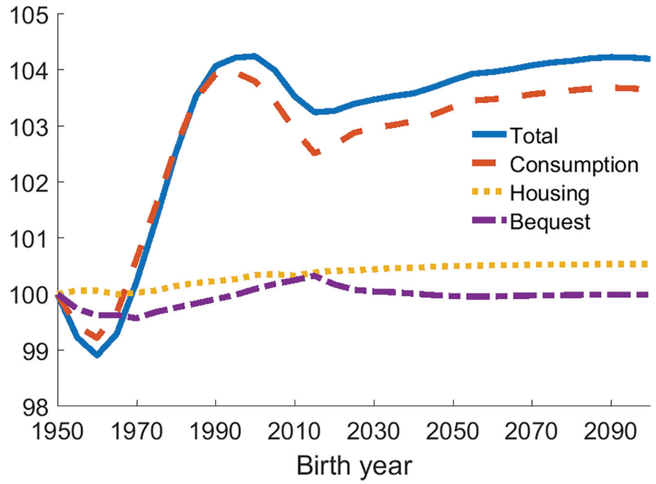
For wealth inequality the trend is more muted, as the two drivers work in opposite directions. The initial increase from 1965 to 1990 as well as the post-2030 increase are mostly due to a more unequal wealth distribution across households, with younger agents being more indebted and middle-aged agents holding higher financial wealth. These higher levels of debt and savings are the flip side of the consumption smoothing described above. Between 1990 and 2030, however, the age composition effect offsets this: the share of poorer, younger households in the total population decreases, while the share of richer, older households increases, thus stabilizing the Gini coefficient for a few decades. The increase in housing wealth inequality implied by the model after 2020 is due to similar factors. The share of households aged above 70 increases strongly between 2020 and 2045. These households tend to hold less housing, as they sell back part of it to finance their retirement consumption. As a consequence, the distribution of housing in the population becomes more unequal, increasing the associated Gini coefficient.

Inequality across Cohorts. In our model, cohorts differ across several dimensions: their size and life expectancy according to exogenous demographic trends, but also consumption, wealth, and housing levels given by general equilibrium outcomes. If simply comparing expected lifetime utility across cohorts, we would find a clear increase over time, meaning that an agent born in 1980 has a higher expected utility than one born in 1950. This is, however, the simple consequence of increased longevity and hence being able to enjoy consumption over a longer period of time, and necessarily depends on underlying assumptions regarding the value of being alive rather than dead.

To obtain a meaningful comparison, it is necessary to disentangle the various components at play, most importantly to abstract from the change in longevity. To do this, we use the decomposition suggested by Jones and Klenow (2016), which allows us to subtract the longevity component of utility from the total utility variation.[21] The result is shown in the blue solid line in figure 13, in consumption

[21] Full details of this exercise are provided in online appendix B.2.

Figure 13. Decomposition of Expected Lifetime Utility (consumption equivalents, %)



Note: Death probabilities fixed at 1950-cohort level.

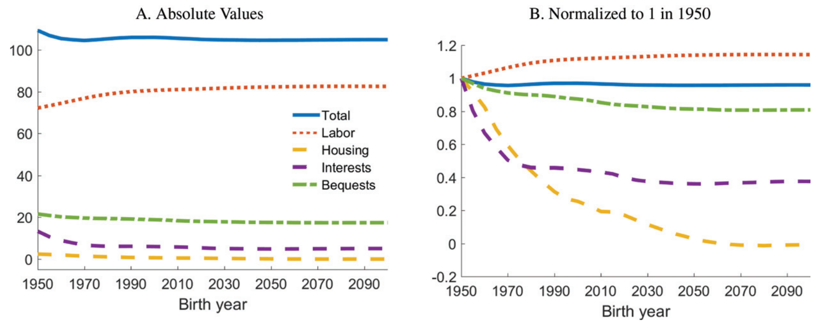
equivalents with respect to the 1950 cohort. This is actually equivalent to comparing utility across cohorts, keeping the death probabilities fixed at the level of the 1950 cohort.[22] Once abstracting from longevity effects, the change in realized utility across cohorts is much smaller and does not monotonically increase, with baby-boomers and households born around 2015 appearing to be worse off.

We can further decompose this total utility across components related to housing, bequests, and consumption in the dashed lines in figure 13.[23] The consumption component is clearly the main driver, with both the magnitude and dynamics looking very similar to total utility. The contribution of housing utility slightly increases for later cohorts, due to a better distribution of housing over the life cycle,

[22]The choice of the cohort here affects the precise magnitude of the utility differentials between cohorts, but does not affect the overall dynamics.

[23]Each component is an aggregate of the level and "smoothing" component in the Jones and Klenow (2016) methodology; see online appendix B.2 for the full decomposition.

Figure 14. Lifetime Income and Its Components



allowed both by the increased ability of young households to borrow and buy housing, and by the changes in the age distribution of the population. The overall contribution of housing utility remains very small, however, as its share in total utility is small (less than 2 percent). Finally, the contribution of utility gained from leaving bequests to the next generations does not show any clear trend over time, and is anyway very small.[24]

Examining income differences across cohorts is another way to understand the origins of their welfare differences. While total income is relatively stable across cohorts, figure 14 shows that this is hiding an increase in labor income and a decrease in capital, housing, and bequests income. Because of the decrease in interest rates and the progressive slowdown in housing prices, income from housing and interest on capital drops dramatically. Their aggregate impact remains small, however, because housing and interest only account for a small share of total income.[25]

[24]The bequests line corresponds to the additional utility from leaving bequests included in the households' utility function, and not to the utility gains obtained by the households receiving the bequests.

[25]Interest income is equal to the net capital share, which is the product of the capital-output ratio, which ranges between 3 to 4, times the annual real interest rate, which ranges between 1 percent and 4 percent. As shown in figure 9, housing wealth increases by around 1 percent of GDP per year from 1950 to 2030 and then flattens off. Accordingly, these capital gains cannot form a large share of any generation's permanent income, though they accrue disproportionately to the early baby-boom generations.

To conclude, increasing longevity implies an increase in welfare, but its impact is difficult to measure. Abstracting from longevity, changes in lifetime consumption are the main drivers of realized utility and seem to be detrimental to the baby-boom generations. The effect of capital gains from housing are lower than one might expect. Finally, an improved consumption smoothing across age, thanks to lower borrowing costs, is found to have a non-negligible positive impact on welfare.

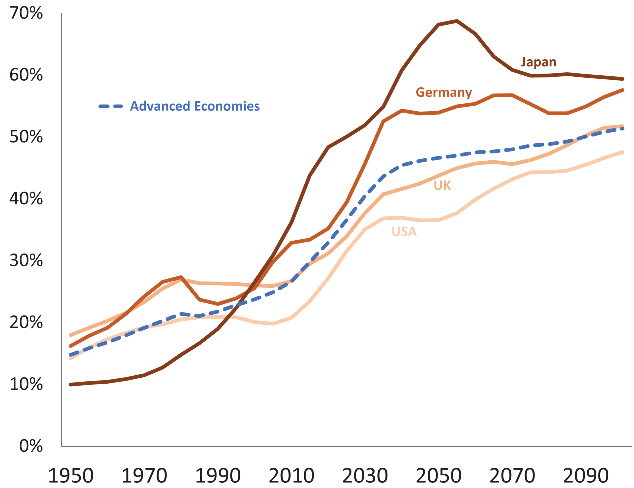
4.2 Open-Economy Implications

So far we have considered the group of advanced economies as one closed economy, and looked at the effects of the demographic trends in the aggregate population. While an aging population is common to all these countries, different countries within this group are aging at different speeds. Figure 15 shows the OADR for a handful of countries within our aggregate group. As can be seen, Japan and Germany, for example, are aging much faster than the aggregate, while the United Kingdom and United States are aging more slowly.

How can our model account for these differences? Consider each of these countries as a small open economy trading on fully integrated global capital markets. In other words, each country takes as given the global real interest rate that arises in the aggregate. All else equal, this means that the firms in each country demand the same level of capital relative to output, which can be seen from their first-order condition. There is, however, no market clearing condition for the domestic capital markets, implying that household savings can be above or below the capital demanded by firms. This discrepancy between domestic savings and domestic capital gives rise to a nonzero net foreign asset (NFA) position for the domestic economy. In particular, if domestic savings are higher than domestic capital, then domestic households must place their savings into capital abroad. Conversely, if domestic capital is higher than domestic savings, some of the domestic capital must be owned by foreign households.

Consider a country such as the United States, which is aging more slowly than the average. There, demographic trends are putting less upward pressure on savings, and hence the global real interest rate is below the interest rate that would arise were the United States a

Figure 15. OADR across Countries



Source: UN Population Statistics (projections based on medium-fertility scenario).

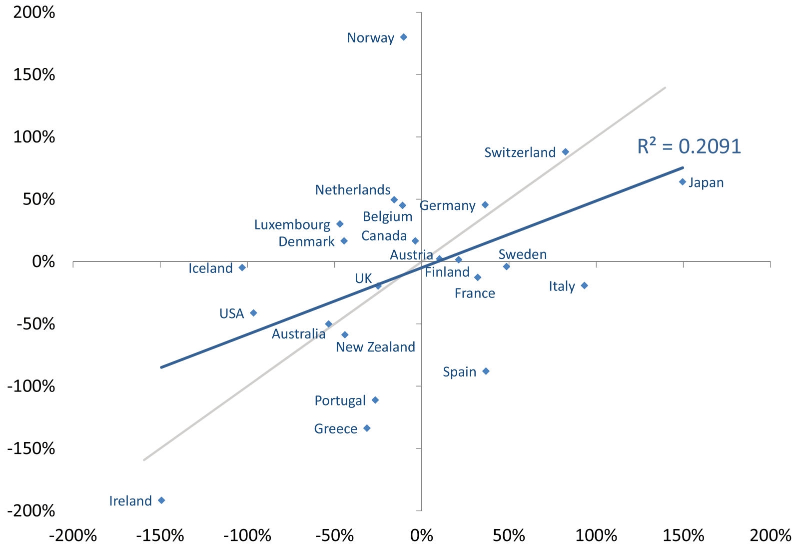
closed economy. In other words, the savings of domestic households in the United States is below the desired capital level of U.S. firms. This translates to a negative NFA position for the United States, as capital flows into the United States from foreign households. Conversely, for a country such as Germany, which is aging faster than the average, the global interest rate is above the rate that would balance the domestic capital market, and this translates to capital outflows from Germany and the accumulation of foreign assets by German households.

To quantify this, we can solve for equilibrium in a small open-economy version of the OLG model, where the interest rate is exogenous and instead of the capital market clearing condition, we have an equation that defines net foreign assets:

$$\widetilde{NFA}_t = \tilde{A}_{t-1} - \tilde{K}_t.$$

We solve this version of the model dynamically with the exogenous path of the real interest rate set as the path of the real interest

Figure 16. NFA-to-GDP in the Model vs. Data (2015)



Source: International Monetary Fund's International Financial Statistics.
Note: Model on x-axis; data on y-axis; gray line is the 45-degree line.

rate from the aggregate exercise, as shown in figure 7, and feeding in the demographic variables of a given country.[26]

Figure 16 plots the level of the NFA-to-GDP ratio in 2015 against the predicted level from the model across all the countries in our advanced-economies group. This exercise can be interpreted as a test of the mechanisms of our model against the data. The model omits any frictions in the international movement of capital, such as capital controls or home bias in portfolio allocations, which were important features of the world economy at least in the early postwar period.[27] Hence we see that the model predicts slightly larger NFA positions

[26]Notice that, so long as we keep the parameters of the model the same for each country, taking the path of the interest rate from the closed-economy exercise is equivalent to solving the model as a multi-country world economy with perfectly integrated capital markets, abstracting from country-specific differences in real interest rates.

[27]The simulations also assume that the economy is always at the dynamic equilibrium, omitting, for example, the major fiscal and physical consequences of the Second World War.

than we observe in the data, with the trend line in this scatter plot being somewhat shallower than the 45-degree line. Nonetheless, a substantial part of the cross-country differences in NFAs can be explained by the model looking only at differences in demographics. This gives us greater confidence about the mechanisms underlying all of the results from our model.[28]

4.3 Robustness and Extensions

Having discussed the various implications of aging population in our baseline model, we now consider some extensions to see the robustness of our results. For the sake of space, all figures are presented in online appendix C.

We begin by comparing the baseline results against a model in which we exclude housing, in order to highlight the role of housing in the baseline (see online appendix C.1). The fall in the interest rate between 1980 and 2100 is slightly larger in the model without housing: there is a 250 bps decline compared with 233 bps in the baseline. This reflects the role of housing as an alternative savings vehicle, which therefore mitigates the decline in interest rates. The rise in the household debt-to-GDP ratio is substantially lower, again reflecting the fact that purchasing housing at a young age is an important driver of higher borrowing in the baseline.

We also investigate the importance of our assumption of exogenously fixed housing supply per capita, by running an alternative exercise where new housing is "produced" from the consumption good, and hence comes out of the resource constraint. This alternative assumption yields an interest rate drop of 224 bps between 1980 and 2100, just 9 bps smaller than the baseline. If anything, this alternative assumption reinforces the role of housing in mitigating the interest rate decline. However, the small differences between this alternative and our baseline results show that the role of housing is overall unaltered by the way it is produced or introduced in the economy, and that this assumption is not important for our results.

[28]See online appendix B.3 for more details on individual countries' NFA paths and model predictions.

Second, we look at the effect of increasing the retirement age (see online appendix C.2). We compare the path of the interest rate in the baseline, in which the retirement age was held constant at age 65, with an alternative simulation in which the retirement age is held constant at age 70, holding other aspects of the calibration fixed. Since we cannot reliably calibrate labor productivity in old age, we carry out each simulation under two different assumptions about old-age productivity. The first, central, case assumes that productivity at age 65–69 is the same as at 60–64. The second, a credible upper bound, assumes that productivity remains at its peak until the end of working life, whether this is 65 or 70.

We find that the potential effect of a higher retirement age, even as high as age 70, is fairly modest in this model, when compared with the effect that demographic changes will have on the interest rate. This is because five years of additional labor income is not enough to offset the rising proportion of time spent in retirement given increasing life expectancy. The extra labor supply is offset in part by more capital, such that the capital-labor ratio, and hence the interest rate, is not affected very much. Even when we assume high productivity in old age, households still face an increasing need to save while they are working in order to finance consumption in the years that they do not have any labor income.

Finally, for comparison with the literature, we use our model to consider the case of the United States more specifically, as a closed economy (see online appendix C.3). The United States has a more dynamic population growth than the average of advanced economies and a life expectancy below average. As expected, we find that the impact of demographic change on the interest rate in the United States is smaller than for the advanced-economy average baseline.

5. Conclusions

In this paper we use an overlapping-generations model, calibrated to advanced-economy data, to assess the contribution of population aging to the fall in real interest rates and other macroeconomic trends which the world has seen over the past three decades. We find that global demographic change can explain three-quarters of the 210 bps fall in global real interest rates since 1980, and larger fractions of the rises in house prices and debt. Importantly, the sign of these effects will not reverse as the baby-boom generations retire: demographic change is forecast to reduce rates by a further 46 bps by 2050. Our model can also explain about 20 percent of the variation of advanced-economy NFA positions.

Among the many uncertainties contained in our analysis, we conclude by highlighting the three most important. The first relates to individual behavior and, in particular, the prediction in our model that households will respond to higher life expectancy with increased saving. How much of these demographic changes are actually anticipated by households in reality? There is limited evidence in the literature showing that savings rise as life expectancy rises. De Nardi, French, and Jones (2009) use variations in life expectancy by gender, initial health, and permanent income to show that higher life expectancy does lead to higher savings, but their focus is on the savings behavior of retirees rather than workers. Similarly, both Bloom, Canning, and Graham (2003) and Kinugasa and Mason (2007) use cross-country panel regressions to show that higher average life expectancy can explain higher national savings rates, but they do not address the potential reverse causation from higher wealth to higher life expectancy due to availability of health care and sanitation.

The second caveat relates to the absence of government pension schemes in our model. The implications of this simplification may vary across countries and actual pension systems. Including a fully funded pension system would preserve the impact of aging on aggregate savings, with pension savings supplied by pension funds instead of being directly supplied by households. Including a pay-as-you-go pension system, by redistributing from workers toward retirees via contemporaneous transfers, could reduce incentives to save in anticipation of retirement. However, the household savings in middle age will still tend to rise in response to higher life expectancy, to the extent that public pensions are insufficient to finance a household's desired retirement. Conversely, maintaining a household's income throughout a longer retirement, by taxing a shrinking working population, would raise sustainability issues, and households could again revert to private savings in anticipation of lower state pension payouts in the future.

The third uncertainty around our results relates to the global economy and, in particular, to the pace and ultimate extent to which emerging markets and low-income countries integrate into world capital markets. These populations have different demographic profiles than advanced economies: they are generally much younger, although emerging markets are set to age rapidly in the coming decades. Their integration into world capital markets, either directly or indirectly through migration into advanced economies, could potentially mitigate the downward pressure on real interest rates from demographic change. On the other hand, if households or institutions in these economies have a higher propensity to save than advanced economies, they could put further downward pressure on real interest rates.

References

Backus, D., T. Cooley, and E. Henriksen. 2014. "Demography and Low-frequency Capital Flows." *Journal of International Economics* 92 (Supplement 1): S94–S102. 36th Annual NBER International Seminar on Macroeconomics.

Barany, Z., N. Coeurdacier, and S. Guibaud. 2018. "Capital Flows in an Aging World." CEPR Discussion Paper No. 13180.

Bloom, D. E., D. Canning, and B. Graham. 2003. "Longevity and Life-cycle Savings." *Scandinavian Journal of Economics* 105 (3): 319–38.

Caballero, R. J., E. Farhi, and P.-O. Gourinchas. 2017. "Rents, Technical Change, and Risk Premia Accounting for Secular Trends in Interest Rates, Returns on Capital, Earning Yields, and Factor Shares." *American Economic Review* 107 (5): 614–20.

Carvalho, C., A. Ferrero, and F. Nechio. 2016. "Demographics and Real Interest Rates: Inspecting the Mechanism." *European Economic Review* 88 (September): 208–26.

De Nardi, M., E. French, and J. B. Jones. 2009. "Life Expectancy and Old Age Savings." *American Economic Review* 99 (2): 110–15.

Domeij, D., and M. Flodén. 2006. "Population Aging and International Capital Flows." *International Economic Review* 47 (3): 1013–32.

Eggertsson, G. B., N. R. Mehrotra, and J. A. Robbins. 2019. "A Model of Secular Stagnation: Theory and Quantitative Evaluation." *American Economic Journal: Macroeconomics* 11 (1): 1–48.

Fernald, J. G. 2012. "A Quarterly, Utilization-adjusted Series on Total Factor Productivity." Working Paper No. 2012-19, Federal Reserve Bank of San Francisco.

Gagnon, E., B. K. Johannsen, and J. D. López-Salido. 2021. "Understanding the New Normal: The Role of Demographics." Forthcoming in *IMF Economic Review*. Available at https://doi.org/10.1057/s41308-021-00138-4.

Goodhart, C., and M. Pradhan. 2017. "Demographics Will Reverse Three Multi-decade Global Trends." BIS Working Paper No. 656.

Holston, K., T. Laubach, and J. C. Williams. 2017. "Measuring the Natural Rate of Interest: International Trends and Determinants." *Journal of International Economics* 108 (Supplement 1): S59–S75.

Jones, C. I., and P. J. Klenow. 2016. "Beyond GDP? Welfare across Countries and Time." *American Economic Review* 106 (9): 2426–57.

Kinugasa, T., and A. Mason. 2007. "Why Countries Become Wealthy: The Effects of Adult Longevity on Saving." *World Development* 35 (1): 1–23.

Kiyotaki, N., A. Michaelides, and K. Nikolov. 2011. "Winners and Losers in Housing Markets." *Journal of Money, Credit and Banking* 43 (2–3): 255–96.

Knoll, K., M. Schularick, and T. Steger. 2017. "No Price Like Home: Global House Prices, 1870-2012." *American Economic Review* 107 (2): 331–53.

Krueger, D., and A. Ludwig. 2007. "On the Consequences of Demographic Change for Rates of Returns to Capital, and the Distribution of Wealth and Welfare." *Journal of Monetary Economics* 54 (1): 49–87.

Lisack, N., R. Sajedi, and G. Thwaites. 2017. "Demographic Trends and the Real Interest Rate." Staff Working Paper No. 701, Bank of England.

Marx, M., B. Mojon, and F. R. Velde. 2018. "Why Have Interest Rates Fallen Far Below the Return on Capital." Working Paper No. 2018-01, Federal Reserve Bank of Chicago.

Piketty, T., and G. Zucman. 2014. "Capital is Back: Wealth-Income Ratios in Rich Countries, 1700-2010." *Quarterly Journal of Economics* 129 (3): 1255–1310.

Rachel, L., and T. D. Smith. 2017. "Are Low Real Interest Rates Here to Stay?" *International Journal of Central Banking* 13 (3): 1–42.

Sajedi, R., and G. Thwaites. 2016. "Why are Real Interest Rates So Low? The Role of the Relative Price of Investment Goods." *IMF Economic Review* 64 (4): 635–59.

Waldron, M., and F. Zampolli. 2010. "Household Debt, House Prices and Consumption in the United Kingdom: A Quantitative Theoretical Analysis." Working Paper No. 379, Bank of England.

Central Bank Communication and Disagreement about the Natural Rate Hypothesis*

Carola Conces Binder
Haverford College

About half of professional forecasters report that they use the natural rate of unemployment (u^*) to forecast. I show that forecasters' reported use of and estimates of u^* are informative about their expectations-formation process, including their use of a Phillips curve. Those who report not using u^* have higher and less anchored inflation expectations, and seem to have found the Federal Reserve's state-based forward guidance less credible. The Federal Open Market Committee (FOMC) publishes participants' projections of longer-run unemployment in the Summary of Economic Projections. I document how and when the FOMC participants have disagreed with each other and with the private sector, discussing possible sources of disagreement and implications for credibility.

JEL Codes: E52, E58, E43, D83, D84.

1. Introduction

In the global financial crisis and Great Recession, with policy rates constrained by the zero lower bound (ZLB), central banks intensified their use of communication as a policy tool (Yellen 2012; Williams 2013b; Blinder 2018). This increase in communication-based monetary policy has been accompanied by greater efforts to understand how different economic agents form beliefs and expectations. Survey data on economic expectations reveal notable heterogeneity in the

*I thank Edward Nelson and participants at the Federal Reserve Board and Federal Reserve Bank of Dallas seminars for useful suggestions. Author e-mail: cbinder1@haverford.edu.

beliefs of consumers, professional forecasters, and central bankers themselves (Mankiw, Reis, and Wolfers 2004; Boero, Smith, and Wallis 2008; Romer and Romer 2008; Patton and Timmermann 2010; Coibion and Gorodnichenko 2012; Andrade and Le Bihan 2013; Binder 2017c). Understanding the sources and nature of this disagreement could have important implications for monetary policy and central bank communication (Coibion and Gorodnichenko 2015; Detmers 2016; Falck, Hoffmann, and Hurtgen 2017).

Patton and Timmermann (2010) argue that disagreement in shorter-horizon expectations mostly reflects differences in private information, while disagreement about longer horizons reflects differences in models. Andrade et al. (2016) show that forecasters in the Blue Chip Financial Forecasts survey disagree even in very long-horizon forecasts for output, inflation, and the federal funds rate, and that this disagreement is time varying. They refer to this long-horizon disagreement as *fundamental disagreement*, as it reflects differing views about slow-moving, unobserved economic fundamentals like potential output, the natural interest rate, and the inflation target. These unobserved fundamentals can be difficult to estimate precisely in real time (Orphanides and Williams 2002; Laubach and Williams 2016; Borio, Disyatat, and Juselius 2017; Holston, Laubach, and Williams 2017).

In this paper, I use data from the Federal Reserve Bank of Philadelphia Survey of Professional Forecasters (SPF) to study forecasters' beliefs and disagreement about the economy in the long run, and related implications for central bank communication. I exploit survey questions that ask forecasters whether they use the natural rate of unemployment (u^*) to make forecasts and, if so, asks for their estimates of u^*. These questions provide explicit and previously underutilized information about forecasters' models and beliefs.

I show that forecasters' responses to these questions are informative about their expectations-formation process. Forecasters who say they use u^* to forecast do appear to do so, in the sense that they expect inflation to fall when they expect unemployment to be above their own estimate of u^*. The inflation expectations of forecasters who report *not* using u^* more closely resemble univariate forecasts and are less sensitive to the unemployment gap or output gap. These results are a novel contribution to the literature on the model consistency of survey forecasts (Pierdzioch, Rulke, and Stadtmann 2011; Rulke 2012; Drager, Lamla, and Pfajfar 2016).[1]

These results also provide empirical support for the general premise of heterogeneous agent models with two types of private agents, distinguished by their expectations formation (Andrade et al. 2018; Beqiraj, Di Bartolomeo, and Di Pietro 2019). In several papers, the two types are "credibility believers" (also called "fundamentalists"), who trust the central bank, expect future inflation to be near the central bank's inflation target, and use a Phillips curve, and "adaptive expectations users" (also called "naive" agents), who use only past inflation to forecast future inflation (Busetti et al. 2017; Goy, Hommes, and Mavromatis 2018; Cornea-Madeira, Hommes, and Massaro 2019; Hommes and Lustenhouwer 2019). Having shown that reported u^* users appear to use a Phillips curve, I next show that they resemble credibility believers in other ways as well. Most notably, their long-run inflation expectations are closer to the Federal Reserve's inflation target and more strongly anchored. Their forecasts are also somewhat more accurate. Thus, while reported use of u^* cannot account for all differences between forecasters, it does seem to provide a useful way to roughly categorize them into these two types.

The presence of credibility believers and adaptive expectations users can have important implications for macroeconomic dynamics and policy. Goy, Hommes, and Mavromatis (2018) study forward guidance at the ZLB in a New Keynesian model with these two types, assuming that only the credibility believers respond to forward guidance.[2] With a smaller share of credibility believers,

[1]Pierdzioch, Rulke, and Stadtmann (2011) find that professional forecasters in the G-7 make forecasts consistent with Okun's law. Rulke (2012) finds that forecasters in Asian-Pacific countries use Okun's law and the Phillips curve. Drager, Lamla, and Pfajfar (2016) show that the share of U.S. consumers and forecasters holding expectations consistent with the Fisher equation, Taylor rule, and Phillips curve is time varying and that central bank communication can facilitate understanding of these rules.

[2]As in Campbell et al. (2012), forward guidance may be Delphic or Odyssean. Delphic forward guidance conveys information about the central bank's outlook, while Odyssean forward guidance is interpreted as a commitment to deviate from the central bank's policy rule in the future, keeping rates "lower for longer" when inflation and growth later rise (Eggertsson and Woodford 2003; Campbell et al. 2019).

forward guidance is less effective. Thus, the presence of adaptive expectations users helps resolve the "forward-guidance puzzle," or the implausibly large responses of macroeconomic variables to forward guidance in standard New Keynesian models with rational expectations (Del Negro, Giannoni, and Patterson 2013; McKay, Nakamura, and Steinsson 2016).[3] To test whether the forecasters who report using u^* resemble credibility believers with respect to forward guidance, I focus on the threshold-based forward guidance issued in December 2012. I find that, indeed, forecasters who report using u^* were less likely to expect liftoff with unemployment above the 6.5 percent threshold announced in the FOMC's forward guidance.

Monetary policymakers communicate not only about the future path of the policy rate but also about their projections of future conditions and estimates of important parameters, including u^*. The quarterly Summary of Economic Projections (SEP) publishes individual FOMC participants'[4] anonymized projections for real gross domestic product (GDP) growth, the unemployment rate, and inflation at several horizons. Longer-run projections for growth, unemployment, and headline inflation were added to the SEP in February 2009, and projections of the longer-run federal funds rate were added in January 2012.[5] The longer-run inflation projections were widely interpreted as an informal inflation target until the January 2012 "Statement on Longer-Run Goals and Monetary Policy Strategy" made the 2 percent inflation target explicit (Orphanides 2019). According to Bernanke (2016b), the longer-run unemployment, output growth, and federal funds rate projections can be interpreted

[3]Other papers that introduce departures from rational expectations, including imperfect knowledge and learning, to attempt to resolve the forward-guidance puzzle include Ferrero and Secchi (2010), Cole (2015), Honkapohja and Mitra (2015), and Eusepi and Preston (2018).

[4]Following Bernanke (2016a), I use "FOMC participants" to refer to the seven Board governors and 12 Reserve Bank presidents who contribute projections to the SEP. "FOMC members" refers to a subset of participants, the seven Board members, the president of the Federal Reserve Bank of New York, and a rotating group of 4 of the remaining 11 Reserve Bank presidents.

[5]https://www.federalreserve.gov/monetarypolicy/timeline-summary-of-economic-projections.htm.

as estimates of u^*, potential output growth (y^*), and the "neutral" federal funds rate (r^*).[6]

Faust (2016) characterizes the SEP as *decentralized* communication, as it reveals the diversity of policymakers' views without clarifying how this diversity will affect committee policy choices. In contrast, *centralized* communication, like the threshold-based forward guidance, clarifies how the FOMC intends to react to incoming information. Faust argues that decentralized communication can potentially lead to cacophony and confusion.[7] For this reason, Bernanke (2016a) judges that the SEP "remains a controversial part of the Fed's communications toolkit, and it has sometimes confused more than enlightened" (also see Thornton 2015, Olson and Wessel 2016, and Bundick and Herriford 2017).

Decentralized communications do not fit neatly into the forward-guidance model of Goy, Hommes, and Mavromatis (2018), who assume that the central bank communicates with perfect precision, though they note that this assumption is not always realistic. The final section of this paper focuses on the FOMC's longer-run unemployment projections, documenting how and when the FOMC participants have disagreed with each other and with the private sector, discussing possible sources of disagreement and implications for credibility.

This paper contributes to several other strands of literature, including strands that use survey measures of expectations to study inflation targeting and expectations anchoring (Davis 2012; Kumar et al. 2015; Binder 2017a), to measure the effects of unconventional monetary policy on private-sector expectations (Bauer and Rudebusch 2013; Swanson and Williams 2014; Engen, Laubach, and Reifschneider 2015; Andrade et al. 2018), or to analyze the nature of information rigidities and the expectations-formation process (Mankiw, Reis, and Wolfers 2004; Coibion and Gorodnichenko 2012). The paper also contributes to a literature on why monetary policymakers disagree and how they communicate disagreement. Nechio and Regan (2016) show that monetary policymakers'

[6] FOMC participants also communicate about these estimates in speeches; e.g., see Clarida (2019).

[7] Papers that suggest that more transparency is not always optimal include Morris and Shin (2002), Thornton (2003), Stasavage (2007), and Sunstein (2017).

speeches reveal a diverse set of views among the FOMC. Hayo and Neuenkirch (2013) and Jung and Latsos (2015) show that regional economic variables affect the interest rate preferences and communications of Federal Reserve presidents.

Finally, the paper contributes to the literature on the natural rate hypothesis (NRH) and its use by policymakers. Friedman (1968) famously argued that there is no long-run tradeoff between inflation and unemployment; rather, unemployment returns to its "natural" rate in the long run. This natural rate is the rate that would be observed when prices and wages have had time to fully adjust to balance supply and demand, and depends on structural factors characterizing the labor market (Walsh 1998).

Blanchard (2018) reviews the arguments and empirical evidence for and against two subhypotheses of the NRH: the *independence subhypothesis*—that there exists a natural rate of unemployment independent of monetary policy—and the *accelerationist subhypothesis*—that monetary policy cannot sustain unemployment below u^* without higher and higher inflation. First, there is some evidence of hysteresis, or path dependence in the natural rate of unemployment, which challenges the independence hypothesis (Ball 2009; Abraham et al. 2019; Yagan 2019). Second, prolonged high unemployment after 2009 did not lead to lower and lower inflation, which challenges the accelerationist hypothesis, though alternative explanations for the "missing disinflation" have been suggested (Coibion and Gorodnichenko 2015). Farmer (2013) also critiques the usefulness of the NRH in explaining inflation dynamics.

Several authors estimate policymakers' beliefs about u^* statistically, via estimation of a model of the economy and of policymakers' learning dynamics (Orphanides and Williams 2005, 2006; Sargent, Williams, and Zha 2006; Williams 2006). Typically, these models include an IS curve and a Phillips curve, written in terms of time-varying natural rates of unemployment and interest that are unobservable to policymakers but follow some specified data-generating process, and a policymaker loss function. Policymakers' misperceptions of u^* can have important implications for inflation dynamics and may have contributed to the Great Inflation of the 1970s (DeLong 1997; Romer and Romer 2002; Reis 2003; Primiceri 2006; Ashley, Tsang, and Verbrugge 2018). Orphanides and Williams (2002) study a variety of generalized Taylor (1993)-type monetary policy rules and show that the most robust rules under such misperceptions are “difference rules” in which the policy rate is raised or lowered from its previous level in response to inflation and changes in economic activity. In contrast to these papers, I use survey-based rather than model-derived measures of policymakers’ beliefs, and examine empirically the heterogeneity in both policymaker and private-sector beliefs.

Others use a narrative approach to study policymakers’ beliefs about u^* and the Phillips curve. For example, Romer and Romer (2004) examine the narrative record to show that Federal Reserve chairs since 1936 have held a variety of views about the sensitivity of inflation to labor market slack and the level of u^*. Meade and Thornton (2012) use FOMC transcripts to evaluate the role of the Phillips-curve framework in U.S. monetary policy from 1979 to 2003. Most policymakers thought that inflation should be related to the gap between aggregate demand and aggregate supply, but disagreed about the usefulness of various gap measures in predicting inflation and guiding policy. I similarly use a narrative approach to supplement my analysis of policymakers’ beliefs. In addition to FOMC transcripts and materials, I also examine the financial and popular press, as my interest is in not only policymakers’ beliefs but also private-sector beliefs.

2. Forecasters’ Use of u^* and Expectations Formation

The Federal Reserve Bank of Philadelphia Survey of Professional Forecasters is a quarterly unbalanced panel of approximately 60 anonymous respondents. I make use of SPF forecasts for the civilian unemployment rate (u), headline PCE inflation (π), and nominal interest rates (i) at multiple horizons. Let $x^{\tau}_{j,t}$ denote forecaster j’s expectation in quarter t of variable x at time τ, where τ may be a calendar year or a quarter depending on context. SPF respondents provide forecasts for the previous quarter (“backcast”), current quarter (“nowcast”), and one, two, three, and four quarters ahead, as well as annual average forecasts for the calendar year in which the survey is conducted and the following calendar year. Beginning in 2009:Q2 and 2009:Q3, respectively, forecasters also provide unemployment and three-month Treasury-bill

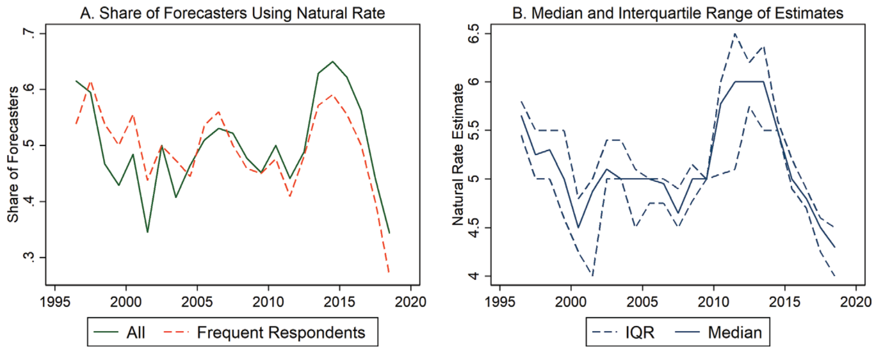
(T-bill) rate forecasts for the subsequent two calendar years. Since 2007:Q1, the SPF collects forecasts of personal consumption expenditures (PCE) inflation for an additional calendar year and averaged over the next five years (from the fourth quarter in the year before the survey year to the fourth quarter of the year that is five years beyond the survey year). SPF T-bill forecasts are for the quarterly or annual average of the underlying daily levels and unemployment forecasts for the average of the underlying monthly levels. Quarterly PCE forecasts refer to annualized quarter-over-quarter percent changes of the quarterly average seasonally adjusted price index, and annual PCE forecasts refer to inflation from the fourth quarter of the previous year to the fourth quarter of the year indicated.

A special SPF segment in 2009 asks respondents about their forecasting methods. Of the 25 forecasters who answered this optional segment, 20 say they use a model with subjective adjustments, 1 uses a model alone, and 4 use just experience and intuition. Of those using a model, 6 say they use a structural model, 3 use univariate or multivariate time-series forecasting, and 11 use some combination. Respondents to the special segment are not identified by forecaster ID, so their reported forecasting methods cannot be matched with their responses to other questions. However, another question on the survey provides information about forecasters' models and methods that can be matched with their forecasts. Namely, in the third quarter of each year since 1996, the SPF asks whether respondents use the natural rate of unemployment (u^*) in forecasting and, if so, asks for their estimate of u^*.

Panel A of figure 1 shows the share of forecasters who report using u^* to forecast over time. The share was around 50 percent during the ZLB period, peaked at 65 percent in 2014, then declined to 34 percent in 2018. While 124 forecasters have responded at least once to the question of whether they use u^*, some of these forecasters have only responded a few times. To address concerns about compositional effects, I also consider the sample of 30 forecasters who have responded to this question in at least 10 years. The share using the natural rate is similar for the full sample and frequent responders.

Panel B shows how the median and interquartile range of estimates of u^* have evolved over time. In 2009:Q3, the 25th and

Figure 1. SPF Forecasters' Use and Estimates of Natural Rate of Unemployment



Notes: Data are from SPF. Frequent respondents are those that provide at least 10 responses to the question of whether they use the natural rate of unemployment.

75th percentile SPF forecasters agreed that u^* was 5 percent. Two years later, the median rose to 6 percent, and disagreement also increased: the 25th percentile was 5.1 percent and the 75th percentile 6.5 percent. The median remained at 6 percent in 2012 and 2013, and fell to a record low of 4.3 percent in 2018:Q3. These estimates are also similar for the frequent responders. Thus, forecasters disagree about whether u^* is a useful forecasting concept, and among those forecasters who do use u^*, there is also time-varying fundamental disagreement about the level of u^*.

2.1 Short-Run Inflation Expectations

What does it mean if a forecaster reports using the natural rate of unemployment to forecast? Recall that according to Blanchard (2018), a key implication of the NRH—typically embedded in a Phillips curve—is that unemployment below u^* will lead to higher inflation. I test whether this implication is observed in forecasters' inflation expectations.

As a baseline, I consider the Phillips curve specification that Williams (2006) uses to study policymakers' beliefs about u^*, which relates inflation (π) to its own lags and the lagged unemployment gap:[8]

$$\pi_t = \gamma_1 \pi_{t-1} + \gamma_2 \pi_{t-2} + \gamma_3 (u_{t-1} - u^*_{t-1}) + \nu_t. \quad (1)$$

For forecasters who provide an estimate $u^*_{j,t}$, I can iterate equation (1) forward one period, apply the expectations operator with respect to forecaster j in quarter t, and estimate the coefficients by regressing her one-quarter-ahead forecast of inflation ($\pi^{t+1}_{j,t}$) on her nowcast and backcast of inflation ($\pi^t_{j,t}$ and $\pi^{t-1}_{j,t}$) and her perception of the unemployment gap ($u^t_{j,t} - u^*_{j,t}$).[9] The first column of table 1 shows that the estimate of γ_3 is negative (-0.14) and statistically significant, as expected. Moreover, it is within the range of estimates that Williams (2006) obtains from rolling regressions using realized data from 1950 to 2003. The median of his rolling regression estimates is -0.23.

In the second column, I include the unemployment gap using the forecaster's own estimate $u^*_{j,t}$ as well as using the Congressional Budget Office (CBO) estimate $u^*_{CBO,t}$. Only the coefficient on $u^t_{j,t} - u^*_{j,t}$ is negative and statistically significant (though of course $u^t_{j,t} - u^*_{CBO,t}$ and $u^t_{j,t} - u^*_{j,t}$ are highly correlated). Thus forecasters do appear to use the estimate of u^* that they personally report.

In columns 3 and 4 I compare the expectations formation of forecasters who claim to use the natural rate with those who claim not to. Since the latter do not provide estimates of u^*, I use $u^t_{j,t} - u^*_{CBO,t}$ as the measure of the unemployment gap in both columns for the sake of comparability. The coefficient on the unemployment gap is less than half the magnitude of that for the forecasters who

[8]Williams (2006) includes several additional lags of inflation and imposes a unity sum on the coefficients; for simplicity I just include two lags with no constraint on the coefficients. Williams's model is a version of the Rudebusch and Svensson (1999) model, but written with a time-varying u^* instead of output gap.

[9]Note that by using nowcasts and backcasts of inflation to estimate equation (1), I avoid the need to make assumptions about forecasters' real-time information about macroeconomic variables, but instead rely on their self-reported knowledge of conditions at time t and $t-1$. This is useful because inflation data are revised frequently, so analysis that assumes that ex post revised data are part of agents' information sets can be misleading (Orphanides 2001).

Table 1. The Natural Rate Hypothesis and SPF Inflation Expectations

	(1) $\boldsymbol{\pi}_{j,t}^{t+1}$	(2) $\boldsymbol{\pi}_{j,t}^{t+1}$	(3) $\boldsymbol{\pi}_{j,t}^{t+1}$	(4) $\boldsymbol{\pi}_{j,t}^{t+1}$	(5) $\boldsymbol{\pi}_{j,t}^{t+1}$	(6) $\boldsymbol{\pi}_{j,t}^{t+1}$	(7) $\boldsymbol{\pi}_{j,t}^{t+1}$	(8) $\boldsymbol{\pi}_{j,t}^{t+1}$	(9) $\boldsymbol{\pi}_{j,t}^{t+1}$
$\pi_{j,t}^{t}$	0.28***	0.26***	0.26***	0.33***	0.24***	0.32***			
	(0.03)	(0.03)	(0.03)	(0.04)	(0.02)	(0.04)			
$\pi_{j,t}^{t-1}$	−0.00	−0.01	−0.02	−0.03	−0.02	−0.03			
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)			
$\pi_{j,t}^{t+4}$							0.25**	0.25**	0.47***
							(0.11)	(0.12)	(0.15)
$u_{j,t}^{t} - u_{j,t}^{*}$	−0.14***	−0.22***					−0.15***		
	(0.02)	(0.06)					(0.03)		
$u_{j,t}^{t} - u_{CBO,t}^{*}$		0.06	−0.13***	−0.05***				−0.13***	−0.07***
		(0.06)	(0.02)	(0.01)				(0.02)	(0.01)
Output Gap					0.13***	0.04**			
					(0.02)	(0.02)			
Constant	1.43***	1.48***	1.52***	1.49***	1.57***	1.51***	1.39***	1.41***	1.11***
	(0.07)	(0.08)	(0.08)	(0.09)	(0.06)	(0.10)	(0.18)	(0.21)	(0.35)
N	827	720	720	679	847	696	822	838	673
R_w^2	0.28	0.25	0.25	0.30	0.26	0.30	0.08	0.10	0.09
R_b^2	0.47	0.43	0.43	0.45	0.32	0.46	0.35	0.31	0.46
Sample	Use u^*	Use u^*	Use u^*	No u^*	Use u^*	No u^*			

Notes: Standard errors are in parentheses. ***, **, and * denote $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively. Time sample is 2007:Q1 through 2018:Q3. Data are from the Survey of Professional Forecasters (SPF). Dependent variable is forecast for next-quarter PCE inflation. $u_{i,t}^*$ is the respondent's estimate of natural rate and $u_{CBO,t}^*$ is the CBO estimate. In columns 4 and 6, sample is the respondents who report that they do not use u^*.

use u^*.[10] Also note that the coefficient on inflation is 0.33, compared with 0.26 for the u^* users. Quarterly PCE inflation has an AR(1) coefficient of 0.30 since 1996 and 0.36 after 2008.

It is possible that respondents who report not using u^* do use a Phillips curve to forecast, but with the output gap in place of the unemployment gap. Columns 5 and 6 are analogous to 3 and 4, but with the output gap in place of the unemployment gap.[11] The coefficient estimate on the output gap for the reported non-users of u^* users is again less than half of that for reported u^* users. Results are robust to alternative specifications of the Phillips curve in (1), including forward-looking specifications. For example, the final three columns are analogous to columns 1, 3, and 4 but include $\pi_{j,t}^{t+4}$ as a regressor in place of $\pi_{j,t}^{t}$ and $\pi_{j,t}^{t-1}$, and results are similar.

In summary, forecasters who report using versus not using u^* appear to be distinct in how they form short-run inflation expectations, and in particular in their beliefs about the Phillips curve. The u^* users seem to rely more on a Phillips curve to forecast short-run inflation, much like the "credibility believers" in the models of Goy, Hommes, and Mavromatis (2018), Cornea-Madeira, Hommes, and Massaro (2019), and others.[12] The non-users do not *perfectly* resemble the "adaptive expectations" or "naive" agents, as their inflation forecasts do rely somewhat on their unemployment forecasts, but their inflation backcasts do explain a much larger share of the variance in their inflation forecasts,[13] so their beliefs can be more reasonably approximated as following a univariate model.

---

[10]This difference is statistically significant. If I instead run the regression with both groups of forecasters, and interact the unemployment gap with a dummy variable indicating that the forecaster uses u^*, the coefficient on the interaction term is negative and statistically significant.

[11]The output gap is defined as $100\frac{y-y^*}{y^*}$, where y^* is the CBO estimate of potential real GDP (Federal Reserve Economic Data (FRED) series GDPPOT) and y is real GDP (FRED series RGDPC1).

[12]In these models, the Phillips curve is specified in terms of marginal cost; since SPF respondents do not provide marginal cost forecasts, I instead use the output or unemployment gap.

[13]Regression of $\pi_{j,t}^{t+1}$ on $\pi_{j,t}^{t-1}$ has an R^2 value about three times higher for non-users than for users.

2.2 Long-Run Inflation Expectations

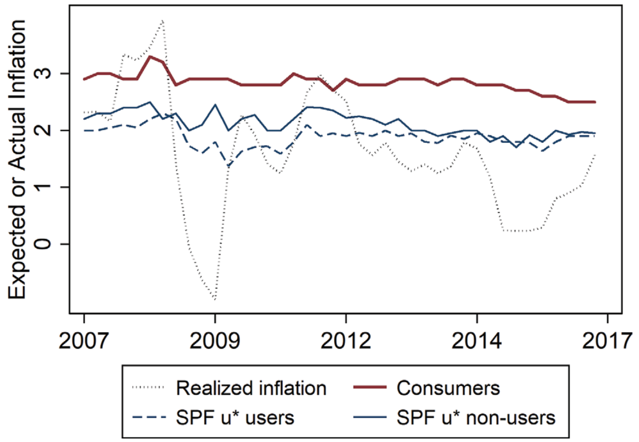
Another important feature of the credibility believers in the models of Cornea-Madeira, Hommes, and Massaro (2019) and others is that they expect future long-run inflation to be equal to the inflation target of the central bank. Busetti et al. (2017) and Hommes and Lustenhouwer (2019) use models with credibility believers and naive agents specifically to study inflation targeting.

January 2012 marked the first explicit announcement of a quantitative inflation target by the Fed, though the Fed had been influenced by the inflation-targeting framework long before this announcement (Bernanke 2003; Thornton 2012). Forecasters who use u^* were more aware that the Fed had an informal inflation target before the 2012 announcement, possibly inferring this from the longer-run inflation projections in the SEP. In a 2007:Q4 special questionnaire, among SPF respondents who have reported at least once that they use u^* in forecasting, 57 percent believed that the Fed had a numerical target for inflation, compared with just 30 percent of other respondents.

Figure 2 plots the median five-year-ahead inflation expectations of each group as well as the longer-run inflation expectations of respondents to the Michigan Survey of Consumers and realized inflation. On average, the u^*-users' long-run inflation expectations are 23 basis points lower than the non-users' expectations, are closer to the inflation target, and fell more in the Great Recession. Non-users' expectations are closer to and more correlated with the consumers' expectations: the correlation between non-users' and consumers' inflation expectations is 0.46, and between users' and consumers' expectations 0.23.

The communication of a numerical target for long-run inflation is intended to make long-run expectations more anchored, or less responsive to shocks. In particular, if expectations are well anchored, long-run inflation expectations should be minimally responsive to changes in shorter-run expectations (Bernanke 2007; Davis 2012). In table 2, I regress forecasters' revisions to five-year-ahead inflation forecasts ($\Delta\pi^{5y}_{j,t} = \pi^{5y}_{j,t} - \pi^{5y}_{j,t-1}$) on revisions to forecasts for the current quarter ($\Delta\pi^{t}_{j,t} = \pi^{t}_{j,t} - \pi^{t}_{j,t-1}$). The sample in the first column is forecasters who report using the natural rate at least once ($NR_j = 1$), and in the second column is forecasters who never report

Figure 2. Long-Run Inflation Expectations of Consumers and Professional Forecasters by Use of u*



Notes: The figure shows median 5- to 10-year inflation expectations of Michigan Survey of Consumers respondents and median 5-year PCE inflation expectations of SPF forecasters who report using or not using the natural rate of unemployment to forecast. Realized inflation refers to the percent change in the PCE price index from one year previous.

using the natural rate ($NR_j = 0$). The natural rate users revise their long-run expectations up 3 basis points for each percentage-point increase in their expectations of current-quarter inflation, compared with 10 basis points for non-users. The R^2 is also much higher for the non-users. Column 3 uses the full sample of forecasters but includes an interaction of $\Delta\pi^t_{j,t}$ and NR_j. The coefficient on the interaction term is negative and statistically significant.

In the fourth column, I consider whether the expectations of the natural rate users became more anchored relative to those of the non-users after the announcement of the inflation target in 2012. This is a diff-in-diff-in-diff specification with the interaction term $Post_t{*}NR_j{*}\Delta\pi^t_{j,t}$, where $Post_t$ denotes that t is after the announcement. The coefficient on the three-way interaction term is negative and statistically significant, suggesting that the announcement may have been more effective at anchoring the expectations of forecasters who use the natural rate.

Table 2. Belief in the Natural Rate Hypothesis and Anchoring of Long-Run Inflation Expectations

	(1) $\Delta\pi^{5y}_{j,t}$	(2) $\Delta\pi^{5y}_{j,t}$	(3) $\Delta\pi^{5y}_{j,t}$	(4) $\Delta\pi^{5y}_{j,t}$
$\Delta\pi^{t}_{j,t}$	0.03***	0.10***	0.10***	0.07***
	(0.01)	(0.02	(0.02)	(0.02)
NR*$\Delta\pi^{t}_{j,t}$			−0.08***	−0.05**
			(0.02)	(0.02)
Post*NR*$\Delta\pi^{t}_{j,t}$				−0.11***
				(0.04)
Post*NR				−0.00
				(0.02)
Post*$\Delta\pi^{t}_{j,t}$				0.12***
				(0.04)
N	1,005	204	1,209	1,209
R^2	0.02	0.13	0.05	0.06
Sample	Use u^*	No u^*	All	All

Notes: Standard errors are in parentheses. ***, **, and * denote $p < 0.01$, $p < 0.05$, and $p < 0.10$, respectively. Time sample is 2007:Q1 through 2018:Q3. Data are from the SPF. Dependent variable is revision in forecast for five-year-ahead PCE inflation. "Post" denotes that the survey date is after the January 2012 inflation target announcement. "NR" denotes that the respondent has reported at least once that she uses the natural rate of unemployment to forecast. Regressions include a constant term and forecaster fixed effects.

2.3 Credibility of Forward Guidance

At the ZLB, central banks' ability to influence private-sector expectations is important; a central bank can conduct monetary easing if it can generate expectations that it will keep the policy rate low to allow above-target inflation and above-trend output in the future (Krugman 1998; Eggertson 2006; Boneva, Harrison, and Waldron 2018). Forward guidance can thus be interpreted as communication about future deviations from the central bank's policy rule (Campbell et al. 2019).

In Goy, Hommes, and Mavromatis (2018), only the credibility believers respond to the central bank's forward guidance. I test whether the reported u^* users resemble credibility believers in this respect. I focus on the "threshold-based" forward guidance of

December 2012 which announced that an "exceptionally low range for the federal funds rate will be appropriate at least as long as the unemployment rate remains above 6-1/2 percent, inflation between one and two years ahead is projected to be no more than a half percentage point above the Committee's 2 percent longer-run goal, and longer-term inflation expectations continue to be well anchored." This guidance was intended to be less ambiguous, and hence better able to guide expectations, than the open-ended guidance issued in December 2008 (Woodford 2012; Williams 2013a).

Swanson and Williams (2014) use multi-horizon forecast data from Blue Chip to infer when private forecasters expected "liftoff" from the ZLB. Before the start of calendar-based forward guidance, the median Blue Chip forecaster expected liftoff in about four quarters. Expected time to liftoff increased in the calendar-based guidance period. Since the Blue Chip data has a maximum horizon of six quarters, for part of the calendar-based forward-guidance period they can only infer that the median forecaster expects liftoff in seven or more quarters. I conduct a similar exercise using the SPF data. Since the SPF data are available at longer horizons, I avoid the top-coding issue and can observe not only the median but also nearly the full distribution of expected liftoff dates. See the appendix for details.

I also compute expected unemployment at expected liftoff for each forecaster and survey date. If an SPF forecaster expects liftoff within the next four quarters, I use her quarterly forecast for unemployment in the corresponding quarter as an estimate of her expected liftoff conditions. If she expects liftoff at a later date, I linearly interpolate between her annual average unemployment forecasts to construct estimates of her expectations of unemployment at each quarterly horizon, and use the interpolated unemployment and inflation forecasts corresponding to my estimate of her expected liftoff date.[14]

In 2013, among forecasters who did not report using u^*, only 33 percent expected unemployment below 6.5 percent at liftoff, compared with 70 percent of forecasters who did report using u^*.

[14] I focus on expected unemployment rather than expected inflation at expected liftoff since the PCE inflation forecasts are available for one less calendar year than the unemployment and T-bill forecasts, and since the 6.5 percent unemployment threshold is clearer that the inflation-related thresholds.

Table 3. Mean Squared Forecast Errors by Use of u^*

	u^{1Q}	u^{4Q}	r^{1Q}	r^{4Q}	π^{1Q}	π^{4Q}
Non-users of u^*	0.15	1.11	0.30	2.17	3.64	3.79
Users of u^*	0.11	0.84	0.22	1.85	4.09	3.61
p-value	0.001	0.01	0.02	0.06	0.50	0.79

Notes: Data are from the SPF. The table shows forecasters' mean squared forecast error by reported use of u^* for unemployment, interest rate (T-bill), and PCE inflation forecasts at the one-quarter and four-quarter horizons. The final row shows the p-value for the test of statistically significant difference in mean between the non-users and users of u^*. Unemployment and interest rate forecast data are available 1996:Q3 to 2018:Q3; inflation forecast data are available 2007:Q1 through 2017:Q3.

Thus, this aspect of forward guidance was more credible among the reported u^* users.

2.4 Forecast Accuracy and Composition of Types

Dragar, Lamla, and Pfajfar (2016) show that "model-consistent" forecasters—those who make forecasts consistent with the Fisher equation, Taylor rule, and Phillips curve—tend to have greater forecast accuracy. I check whether forecasters who report using u^* likewise make more accurate forecasts. Table 3 reports the mean squared forecast error for unemployment, nominal interest rate, and inflation forecasts at the one-quarter-ahead and four-quarter-ahead horizons by reported use of u^*. The forecasters who report not using u^* have larger forecast errors, on average, for unemployment and interest rates at both horizons. The difference in accuracy is statistically significant for unemployment at both horizons and for interest rates at the one-quarter horizon, and marginally significant (p-value $= 0.06$) for interest rates at the four-quarter horizon. The average difference in inflation forecast accuracy is not statistically significant.

The models with credibility believers and naive agents make no assumptions about which type makes more accurate forecasts. Rather they assume, as in Brock and Hommes (1997) and Branch et al. (2004), that agents switch heuristics based on "relative fitness," or some history of relative forecasting performance.[15] That

[15] This assumption formalizes Simon's (1984) suggestion that decisionmaking can be modeled as a rational choice between a set of different heuristics.

is, if the forecasts made by credibility believers become relatively less accurate than the adaptive expectations forecasts, then a larger share of agents will use adaptive expectations.

In Cornea-Madeira, Hommes, and Massaro (2019), agents switch between being credibility believers and adaptive expectations users based strictly on the relative inflation-forecasting performance of the two heuristics. Cornea-Madeira, Hommes, and Massaro estimate the share of credibility believers over time using aggregate data (without survey data on expectations) and a New Keynesian model. They find that the average share of credibility believers has declined in recent years, and posit that in the aftermath of the financial crisis, prolonged below-target inflation has improved the relative forecast accuracy of simple univariate ("naive") forecasts, reducing the share of credibility believers. (See Busetti et al. 2017 for a similar discussion.)

Recall from panel A of figure 1 that the share of u^* users has also declined in recent years. The share of u^* users, which has mean 0.5 and standard deviation 0.08, is not as volatile as the estimated share of credibility believers in Cornea-Madeira, Hommes, and Massaro (2019), which has mean 0.33 and standard deviation 0.27. Part of the difference in mean and volatility may reflect the difference in sample periods, as the sample in Cornea-Madeira, Hommes, and Massaro (2019) starts in 1964 and mine starts in 1996. But it is also possible that forecasters' choice of model depends on more than just inflation forecast accuracy. Forecasters may consider the accuracy of forecasts for multiple variables, or type may be "sticky" due to switching costs (cognitive or otherwise). Forecasters may also evaluate the relative ease of using different models. For example, if u^* becomes highly variable and difficult to precisely estimate, they may switch away from using models that rely on u^*.[16] Forecasters may also be influenced by central bank communications or media narratives about how the economy works.

[16]The absolute number of forecasters that switch from reportedly using to not using the natural rate or vice versa is fairly small: since 1997, only 39 forecasters have switched from not using to using, and only 34 have switched from using to not using. Thus it is difficult to test statistically for possible predictors of switching behavior.

3. Federal Reserve Communication and Disagreement about u^*

The previous section showed that forecasters who report using versus not using the natural rate of unemployment are distinct in how they forecast short- and long-run inflation. The u^* users seem to resemble "credibility believers," including with respect to forward guidance at the ZLB. Thus the time-varying share of u^* users may have important implications for central bank credibility and expectations formation. But recall from figure 1 that even among reported u^* users, estimates of u^* and disagreement about u^* are also time varying. These variations are worth understanding for several reasons.

First, section 2.1 showed that forecasters use their *own* estimates of u^* to form inflation forecasts. The negative estimate of γ_3 in equation (1) implies that, all else equal, forecasters with higher estimates of u^* should have higher expectations of future inflation. Thus disagreement about u^* contributes to disagreement in inflation expectations. This is also true for longer-run inflation expectations. Panel regressions of five-year-ahead inflation expectations on $u^*_{j,t}$ with time fixed effects have a coefficient estimate of 0.23 on $u^*_{j,t}$, which is statistically significant with $p < 0.05$.[17]

Second, the quarterly Summary of Economic Projections publishes FOMC participants' estimates of u^*. The SEP is a decentralized form of Fed communication (Faust 2016). Substantial disagreement about u^* *among* SPF forecasters, or *between* SPF forecasters and FOMC participants, despite publication of the SEP, might point to weaknesses in Federal Reserve communication, and might be related to the subsequent reduction in reported u^* users. As I will show, both types of disagreement were especially high from 2011 through 2013, when many SPF forecasters became more pessimistic than many FOMC participants about u^*. Third, and relatedly, recall from section 2.3 that in 2013:Q3, 33 percent of u^* non-users and 70 percent of u^* users expected unemployment below 6.5 percent at liftoff. Among forecasters with an estimate of u^* less than 6 percent (the highest FOMC projection) in that quarter, 83 percent expected

[17] If the regression includes forecaster fixed effects, the coefficient is 0.19, which is statistically significant with $p < 0.01$.

unemployment below 6.5 percent at liftoff. For those with an estimate of u^* at least 6 percent, only 50 percent expected unemployment below 6.5 percent at liftoff.[18] Thus forecasters who were more pessimistic about u^* than most of the FOMC were less likely to have expectations consistent with the threshold-based forward guidance.

3.1 FOMC Projections of Longer-Run Unemployment

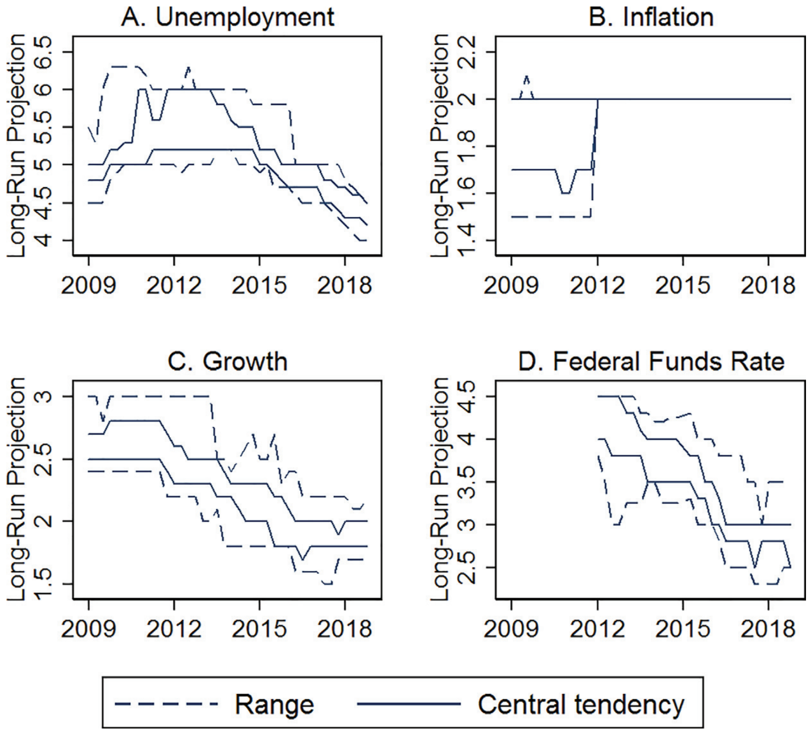
In the SEP, the five Board members and 12 presidents provide projections for several macroeconomic variables for the current calendar year and up to three subsequent years, as well as for the "long run." The projections are not unconditional expectations, but are conditional on appropriate monetary policy. Responses are anonymized and cannot be linked from one meeting to the next.

Panel A of figure 3 summarizes FOMC projections of longer-run unemployment from the SEP, which are available since 2009. In 2009, the FOMC projections in 2009 displayed minimal disagreement, with the central tendency from 4.8 to 5 percent. The width of the central tendency of the FOMC projections subsequently rose, and the midpoint of the central tendency increased. Since projections are conditional on appropriate monetary policy, the increasing width of the central tendency could reflect growing divergence in assumptions about appropriate policy.

These patterns are similar to those for the SPF: panel B of figure 1 shows that in 2009:Q3, the majority of SPF respondents who reported using u^* also estimated that u^* was 5 percent. In fact, for forecasters that said they used u^*, 57 percent reported an estimate of 5 percent, and all estimates were between 4 percent and 6 percent. The median estimate and the interquartile range (disagreement) both rose the next year and remained elevated throughout the ZLB period. But most FOMC projections increased by less than most SPF estimates of u^*. As figure 4 shows, by 2011 through 2013, the SPF median was around 50 basis points higher than the FOMC midpoint. In 2013:Q3, the central tendency of the FOMC long-run unemployment projections was 5.2 to 5.8 percent, and 55 percent of SPF estimates of u^* were *above* 5.8 percent.

[18] This difference is statistically significant at the 10 percent level.

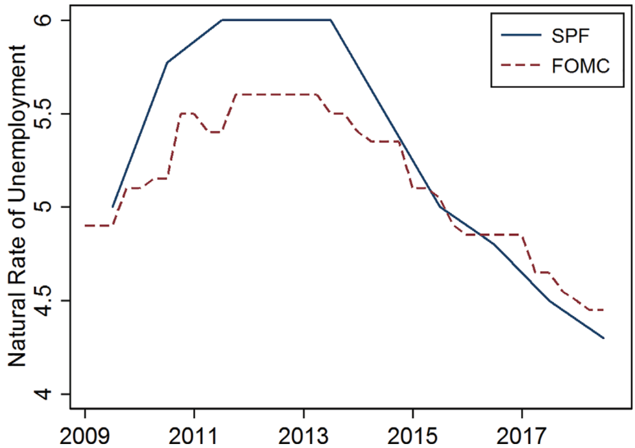
Figure 3. FOMC Longer-Run Projections



Notes: Summary of Economic Projections data accessed from FRED. The central tendency excludes the three lowest and three highest projections. Variable codes: UNRATERLLR, UNRATECTLLR, UNRATECTHLR, PCECTPIRHLR, PCECTPIRLLR, PCECTPICTLLR, PCECTPICTHLR, PCECTPIRHLR, GDPC1RHLR, GDPC1RLLR, GDPC1CTLLR, GDPC1CTHLR, GDPC1RHLR, FEDTARRHLR, FEDTARRLLR, FEDTARCTLLR, FEDTARCTHLR, and FEDTARRHLR.

The other panels of figure 3 summarize FOMC longer-run projections of PCE inflation, growth, and the federal funds rate, while figure 5 shows the width of the central tendency and the range for each longer-run projection over time. Notice that there is no disagreement about longer-run inflation since the 2012 announcement of a 2 percent target. Disagreement about longer-run growth did not increase with disagreement about longer-run unemployment, but rather stayed nearly constant as the midpoint longer-run growth

Figure 4. SPF Estimates and FOMC Projections of Natural Rate of Unemployment



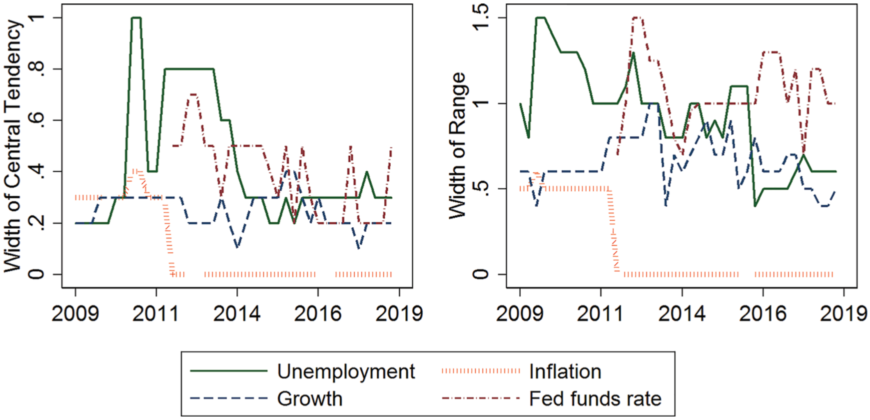
Notes: The solid line is the median SPF estimate of u*. The dashed line is the midpoint of the central tendency for the SEP longer-run unemployment projection.

estimate gradually declined. The longer-run federal funds rate projections, published since 2012, show substantial disagreement about the longer-run policy rate as the midpoint estimate has fallen, which may reflect the documented low precision in estimates of the natural interest rate (Laubach and Williams 2016).

3.2 Definitions of u^*

Why did FOMC and SPF estimates of u^*, which were so similar in 2009, subsequently diverge? One possibility is that forecasters and FOMC participants use different definitions of "natural rate of unemployment." Bernanke (2016b) says that the longer-run unemployment projections in the SEP "can be viewed as estimates of the 'natural' rate of unemployment, the rate of unemployment that can be sustained in the long run without generating inflationary or deflationary pressures." The SPF respondents are not provided with a definition of "natural rate of unemployment."

Figure 5. FOMC Disagreement in Longer-Run Projections

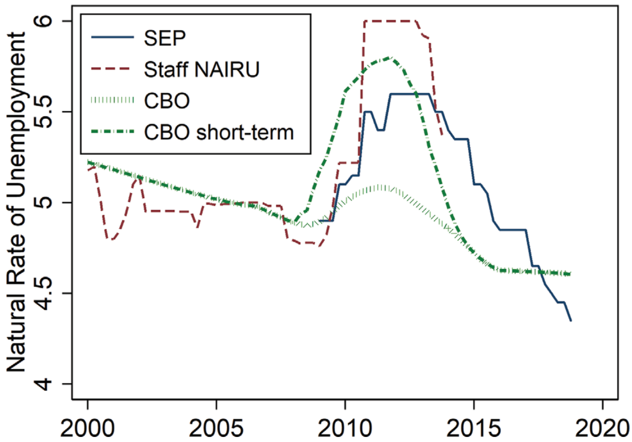


Notes: Summary of Economic Projections data accessed from FRED. The central tendency excludes the three lowest and three highest projections. Variable codes: UNRATERLLR, UNRATECTLLR, UNRATECTHLR, PCECTPIRHLR, PCECTPIRLLR, PCECTPICTLLR, PCECTPICTHLR, PCECTPIRHLR, GDPC1RHLR, GDPC1RLLR, GDPC1CTLLR, GDPC1CTHLR, GDPC1RHLR, FEDTARRHLR, FEDTARRLLR, FEDTARCTLLR, FEDTARCTHLR, and FEDTARRHLR.

The natural rate of unemployment is often treated as synonymous with the NAIRU (non-accelerating inflation rate of unemployment), though the two concepts are distinct and play different roles in monetary policy (Estrella and Mishkin 1999). The NAIRU is the unemployment rate consistent with steady inflation in the near term, and thus plays a more direct role in policy conduct because it helps with forecasting inflation and achieving an inflation target. However, its high variability and difficulty to measure (Staiger, Stock, and Watson 1997; Tasci and Verbrugge 2014) can make the NAIRU problematic to use when explaining policy actions to the public. See Espinosa-Vega and Russell (1997) for a detailed history of economic thought surrounding the NAIRU and the natural rate hypothesis. Meanwhile the natural rate, which is slower moving, serves as the appropriate benchmark for unemployment stabilization objectives (Walsh 1998).

The distinction between the natural rate and the NAIRU may have been minimal in 2009 but larger in the subsequent years.

Figure 6. SEP, Greenbook, and CBO Estimates of Natural Rate of Unemployment or NAIRU



Notes: The solid line is the midpoint of the central tendency for the SEP longer-run unemployment projection. The dashed line is the real-time NAIRU estimate from the Board of Governors, accessed from the Federal Reserve Bank of Philadelphia Greenbook Data Sets. CBO estimates of u^* are accessed from FRED (series NROU and NROUST).

Indeed, until recently, the CBO published a single series they referred to as the "natural rate of unemployment (NAIRU)" (implicitly treating the natural rate and NAIRU as synonyms). But in 2008, the CBO began distinguishing between a "long-run natural rate" and a "short-run natural rate." The latter, which incorporates temporary factors, is more akin to the NAIRU in that it is used to gauge labor market slack in the CBO projections of inflation.

Figure 6 displays both of these CBO series over time. In 2009, the longer-term and shorter-term CBO estimates were 4.9 percent and 5.2 percent, respectively. But the longer-term estimate remained near 5 percent, while the shorter-term estimate peaked at 5.8 percent in 2011:Q4. Figure 6 also plots the midpoint of the SEP longer-run unemployment, which rose much more than the longer-term CBO estimate but less than the shorter-term CBO estimate.

The Federal Reserve Bank of Philadelphia Real-Time Data Research Center provides the "NAIRU Estimates from the Board of

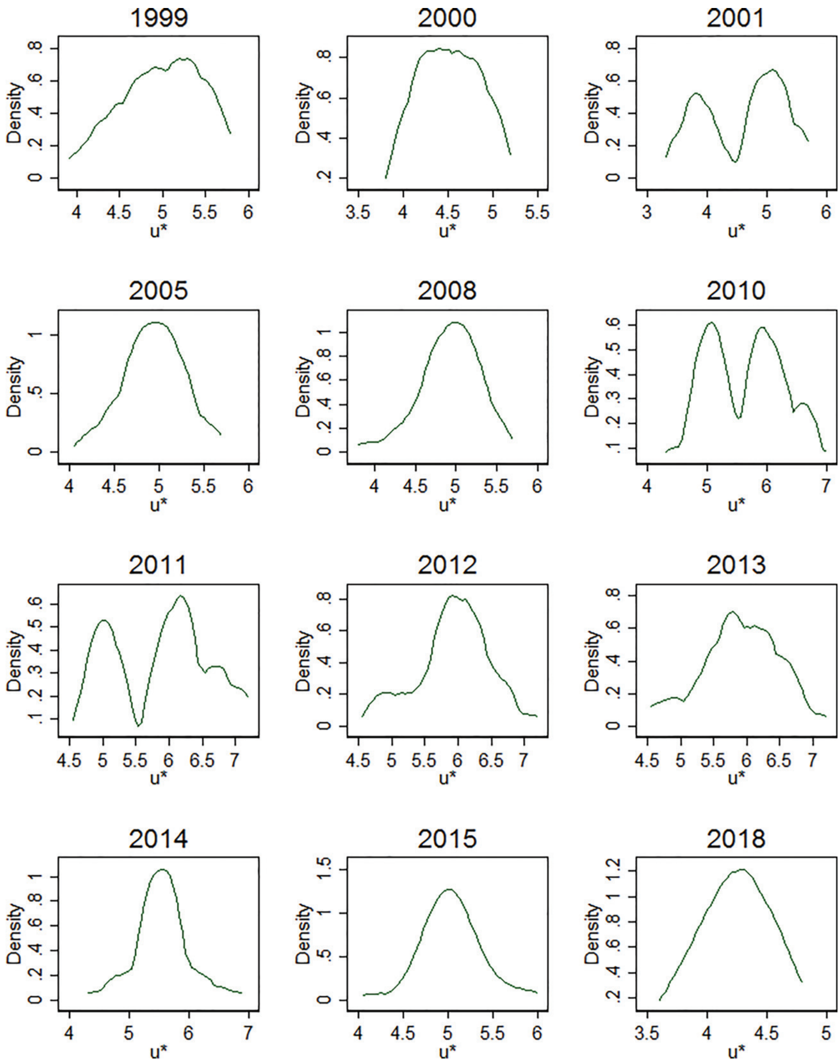
Governors," which contains the Federal Reserve staff's real-time estimates of the NAIRU from the Greenbooks.[19] The data are released with a lag of at least five years; as of August 2019, the NAIRU data are available through December 2013. This NAIRU estimate is also plotted on figure 6. The staff NAIRU estimates rose from 5.0 percent in 2009:Q3 to 6.0 percent in 2010:Q4, and remained at 6.0 percent (well above the SEP midpoint but similar to the SPF median) through 2012:Q4. By 2013:Q3, both the staff NAIRU estimate and the SEP longer-run unemployment midpoint were 5.5 percent. Thus, the SEP longer-run unemployment projections appear to be similar to the staff NAIRU estimates in normal times, but the NAIRU is a shorter-run concept that may rise more than the natural rate of unemployment in recessions.

It is possible that some SPF respondents report estimates of the (short-run) NAIRU while others report the (long-run) natural rate of unemployment. Figure 7 plots kernel density estimates of SPF u^* estimates in different years. The distribution of u^* estimates is unimodal in all years (including the years not displayed) except for 2001, 2010, and 2011, when it is clearly bimodal. In 2001, a recession year with a sharp rise in unemployment, the modes are at 4 percent and 5 percent, while in 2010 and 2011, the two modes are near 5 percent and 6 percent, perhaps corresponding to a group of respondents reporting the natural rate and another group reporting the NAIRU. By 2012 and 2013, though the kernel density appears unimodal, the popular responses are 5.5 percent, 6 percent, and 6.5 percent. The lower estimates may still correspond to respondents reporting the natural rate and higher estimates the NAIRU. This could explain why disagreement among SPF forecasters and between the SPF and the FOMC were both heightened in 2010 through 2013.

However, even among the FOMC participants, disagreement about longer-run unemployment was especially high from 2010 through 2013. Moreover, as FOMC and SPF estimates of u^* have been repeatedly revised downward, the share of reported u^* users

[19] Data are available at https://www.philadelphiafed.org/surveys-and-data/real-time-data-research. For comparative analysis of staff and FOMC members' forecasts, see Romer and Romer (2008) and Binder and Wetzel (2018).

Figure 7. Kernel Density Estimates by Year of SPF Estimates of $u*$



Notes: Data are from SPF. Kernel density estimates for forecasters' estimate of the natural rate of unemployment by year for select years. Epanechnikov kernel with bandwidth 0.2.

has also fallen. The final subsection discusses other potential contributors to disagreement and uncertainty about the natural rate based on narrative evidence.

3.3 Narrative Evidence and Discussion

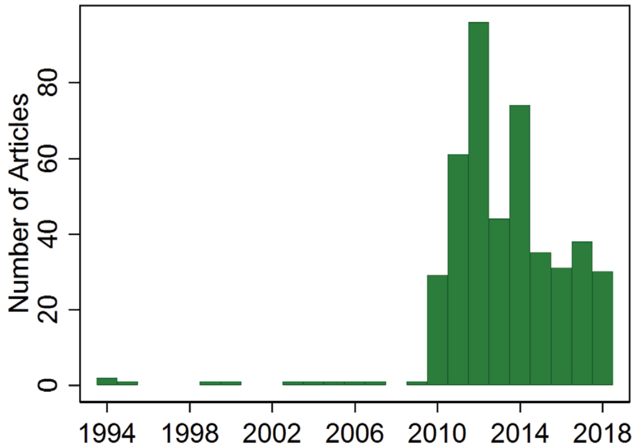
FOMC transcripts and speeches, media coverage, and the academic literature provide some additional insights into the patterns that appear in figures 1, 3, and 4. The rise in median or midpoint estimates and disagreement about u^* for both the SPF and the FOMC corresponds to the timing of the "missing disinflation" puzzle. This puzzle refers to the fact that inflation fell relatively little despite sustained high unemployment in the aftermath of the Great Recession. This missing disinflation led to uncertainty and disagreement about whether the Phillips curve was "alive and well," and about the extent to which a rise in u^* was the cause (Coibion and Gorodnichenko 2015).

Abraham (2015) notes that the idea that the labor market is suffering from "skills mismatch" often becomes popular during prolonged periods of high unemployment. This does appear to be the case following the Great Recession. Paul Krugman describes a consensus by the news media that the high unemployment during and after the Great Recession was structural, resulting from skills mismatch.[20] He argues that the media presented the skills mismatch story as the known truth, despite weak evidence to support it. I searched U.S. publications in the Nexis Uni database for the terms "skill mismatch" or "skills mismatch" and "unemployment." As shown in figure 8, the volume of news coverage of skill mismatch did indeed rise dramatically beginning in 2010 and peaking in 2012.

Reports of skill mismatch often accompanied discussions of the unconventional policies introduced by the Fed at the ZLB, including the quantitative easing (QE) programs (see Blinder 2010). Some drew the conclusion that monetary policy, particularly QE3, would have limited ability to reduce unemployment. For example, on Bloomberg TV, John Ryding, Chief Economist and Founding Partner at RDQ Economics, said, "Let's remember that there's certain

[20] Paul Krugman, "Structural Unemployment: Yes, It Was Humbug," *New York Times*, August 4, 2017.

Figure 8. News Coverage of Skill Mismatch



Notes: Number of search results in U.S. publications in Nexis Uni database for terms "skills mismatch" or "skills mismatch" and "unemployment." There are 450 results from 1994 to 2018. Results can be viewed at https://advance.lexis.com/api/permalink/03d4429d-c5b8-4af6-8eb4-61062b9ed157/?context=1516831.

things that monetary policy can't achieve. ... We have a relatively high level of job openings given the unemployment rate. So why aren't we filling those job vacancies at a faster pace? Is it all about lack of demand, or is it about a skill mismatch?"[21] Some news coverage around the announcement of QE3 also directly or indirectly criticized the Fed's model of the economy. The September 19, 2012 edition of *Forbes* said, "The mistaken belief that a central bank can increase employment is the result of ... persistent theoretical errors. ... One is the so-called 'Phillips curve'. . ."[22]

Forecasters may have differed in the extent to which they believed the skills mismatch narrative, and thus in how much they revised their u^* estimates. FOMC participants, likewise, may have differed in these respects. FOMC transcripts show that as early as 2010,

[21]Source: Interview by Tom Keene, Sara Eisen, and Scarlet Fu on July 25, 2012.

[22]https://www.forbes.com/sites/markhendrickson/2012/09/19/the-wages-of-bernanke-winners-and-losers-from-qe3.

FOMC participants were highly attentive to the possibility of rising structural unemployment, and were relying on their staffs to evaluate this possibility. For example, at the January 2010 FOMC meeting, Janet Yellen, then president of the Federal Reserve Bank of San Francisco, noted that her forecasts were “strikingly optimistic relative to most private-sector forecasters.”[23] Yellen also noted that the median FOMC forecast for 2009 unemployment was 1.5 percentage points too low, explaining that Okun’s law would have predicted 8 percent (rather than the realized 10 percent) unemployment. But her staff could not find evidence that the NAIRU had jumped enough to reconcile the 2009 output and unemployment data with Okun’s law. Instead, she attributed the deviation from Okun’s law to a surge in labor productivity, implying that slack in the economy was higher than in the Greenbook estimates.

At the June 2010 meeting, Chairman Bernanke remarked, “I am still sympathetic to the staff view that the NAIRU—or the natural rate, or however you want to describe it—has probably not permanently increased at this point.”[24] He recommended research by economists at the Federal Reserve Banks of San Francisco and New York, finding that the rate of outflow from unemployment was not industry dependent and was much higher than experienced in the European hysteresis episode in the 1980s (Elsby, Hobijn, and Sahin 2010). Board member Daniel Tarullo also cited Elsby, Hobijn, and Sahin (2010) when he concluded that there were few signs of increasing structural unemployment, and that the severity of longer-term unemployment problems could be limited.[25] Note that FOMC transcripts are released with a five-year lag, but participants could discuss their interpretations of economic conditions in speeches.

In 2012, additional research from Federal Reserve economists continued to find little evidence of skill mismatch or rising structural unemployment (Altig 2012; Fabermand and Mazumder 2012). Jeffrey Lacker, president of the Federal Reserve Bank of Richmond and the only dissenter to the 2012 FOMC statements, viewed structural unemployment and skill mismatch as a larger problem than

[23] January 27, 2010 FOMC transcript, pages 137–39.
[24] June 23, 2010 FOMC transcript, pages 124–25.
[25] April 27–28, 2010 FOMC transcript, pages 130–31.

most other FOMC participants. Lacker publicly remarked that "the healing in the labor market has been limited by a very real sense of skill mismatch," and that unemployment "could well be above 7%" when the Fed raised rates.[26] He made similar comments in speeches throughout the year. For example, in October 2012 he was quoted in Reuters saying that "improvement in labor market conditions appears to have been held back by real impediments that are beyond the capacity of monetary policy to offset."[27] Note that Lacker gave 28 speeches in 2012 and 2013,[28] the same as the average among voting presidents on the FOMC. But he received 22 percent more news coverage than the average voting president and 47 percent more than the average Board member (table 4), as media coverage, including monetary-policy-related coverage, tends to focus on negative news and stories of division and conflict (Binder 2017b).

More recently, inflation has risen relatively little as unemployment has fallen; in other words, there is a "missing inflation" puzzle (Bobeica and Jarocinski 2019). FOMC participants have revised their u^* estimates downward, while SPF forecasters have either revised their estimates downward or stopped using u^* altogether. The repeated downward revisions and "missing inflation" may have contributed to skepticism about the usefulness of models based on the natural rate, reflected in the declining share of SPF respondents who reportedly use u^* to forecast. This skepticism may also extend to some politicians. During Chairman Jerome Powell's July 2019 Congressional testimony, Representative Alexandria Ocasio-Cortez criticized the Fed for its repeated downward revisions to the natural rate estimates and the failure of higher inflation to materialize.[29]

---

[26] May 1, 2012. Bloomberg Panel Word from the Federal Reserve Bank of Richmond. SEC Wire. Retrieved from Nexis Uni.

[27] https://www.reuters.com/article/us-usa-fed-lacker/feds-lacker-says-inflation-fears-underpin-policy-dissent-idUSBRE89P0IF20121026.

[28] Federal Reserve Bank of Richmond Speeches Archive at https://www.richmondfed.org/press_room/speeches/jeffrey_m_lacker?mode=archive.

[29] Her comments were widely publicized, e.g., see http://nymag.com/intelligencer/2019/07/aoc-is-making-monetary-policy-cool-and-political-again.html.

Table 4. News Coverage of FOMC Members

Name	Position	News Items 2012–13
Ben S. Bernanke	Chair	144,442
Janet L. Yellen	Vice Chair	4,514
William C. Dudley	New York	815
Jeffrey M. Lacker	Richmond	812
Dennis P. Lockhart	Atlanta	730
Sandra Pianalto	Cleveland	381
John C. Williams	San Francisco	606
Elizabeth Duke	Board	247
Jerome H. Powell	Board	442
Sarah Bloom Raskin	Board	490
Jeremy C. Stein*	Board	651
Daniel K. Tarullo	Board	942
President Average		669
Board Average (Excluding Chair and Vice Chair)		554

Notes: Data collected by the author from Nexis Uni. I searched for articles including the name (with or without middle initial) and the terms "Federal Reserve" or "FOMC": e.g., ("Jeffrey M. Lacker" OR "Jeffrey Lacker") AND ("Federal Reserve" OR "FOMC"). I limit the date range to January 1, 2012 through December 31, 2013 and the location to the United States.
*Stein joined the FOMC in May 2012.

4. Conclusion

This paper has shown that forecasters who report that they use the natural rate of unemployment to make forecasts are different from forecasters who report that they do not in several key ways. Specifically, the reported users and non-users of u^* seem to resemble the "credibility believers" and "adaptive expectations users," respectively, in the models of Goy, Hommes, and Mavromatis (2018) and others. Forecasters who report using u^* do appear to believe more strongly in the accelerationist hypothesis than those who report not using it, and their reported estimates of u^* are meaningful in the sense that they help predict their forecasts of unemployment and inflation in a model-consistent way. They have longer-run inflation forecasts that are closer to the Fed's target and more strongly anchored, in the sense of being less responsive to changes in short-run inflation expectations. Users of u^* also seem more likely to find forward guidance credible. In other words, I show that differences in forecasters' *self-reported* characterizations of the models they use contribute to forecast disagreement.

The recent large decline in the share of forecasters who report using u^* could potentially increase the communication challenges faced by the Fed. More generally, central bank efforts to influence private-sector expectations may be complicated if private forecasters disagree with the central bank about the model of the economy or key model parameters. Central bankers communicate about their models and forecasts in addition to communicating about policy decision and goals. The Federal Reserve has experimented in recent years with a variety of communication tools, both collective and decentralized. This paper has explored the disagreement about longer-run unemployment reflected in the the Summary of Economic Projections, and how this decentralized communication may have interacted with private-sector beliefs about u^*.

Powell (2016) has noted that "too many voices saying too many different things" contributes to a "cacophony problem" in central bank communication. Future research should consider how central banks can more effectively communicate the models and assumptions underlying their projections to the public, and examine the optimal level and type of transparency when market participants and committee members disagree about fundamentals.

Appendix. Inferring Expected Liftoff Timing

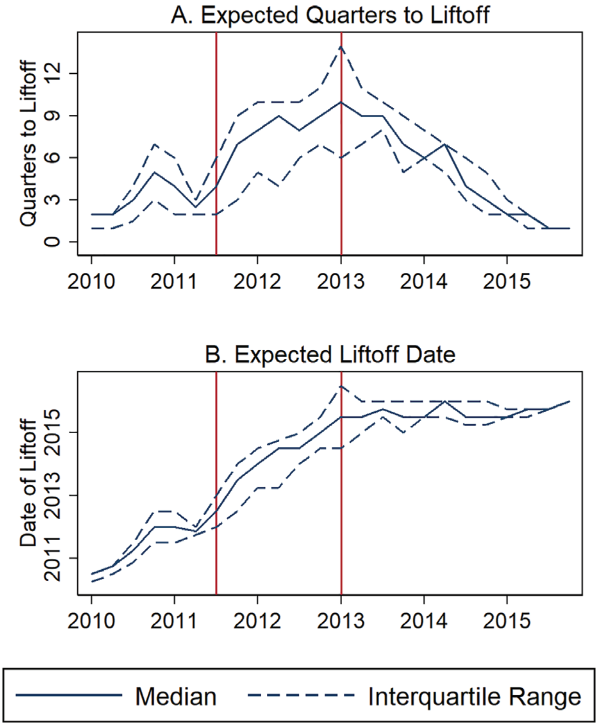
I use multi-horizon nominal interest rate forecast data from the SPF to infer expectations about the timing of liftoff from the ZLB. Since SPF respondents provide quarterly average and annual average instead of year-end forecasts, it is somewhat complicated to infer whether they expect liftoff by the end of a particular calendar year. Suppose that in quarter t, a forecaster expects the interest rate hike of 25 basis points to occur in quarter $t+1$. Each quarter is approximately 13 weeks long, with an FOMC meeting in the third and ninth week. So if she expects the rate hike in the second meeting of the quarter, her quarterly average forecast for the federal funds rate in quarter $t+1$ will be approximately 4 basis points above the ZLB rate. Since the T-bill rate is the average of the expected federal funds rate over the next 91 days, she should expect the T-bill rate to rise a very small amount in quarter t and by around 14 basis points in quarter $t+1$. The difference between her forecast for quarter $t+1$ and her backcast for quarter $t-1$ will be around 14 basis points. If she expects a rate hike in the first meeting, or at both meetings, the difference will of course be greater. With the calendar-year forecasts, an even smaller increase in the T-bill forecast can potentially indicate expected liftoff: by similar reasoning, if she expects liftoff by the last meeting of the calendar year, the annual forecast should be at least 4 basis points above the backcast.

Since a small difference of a few basis points between the backcast and forecast could also indicate uncertainty (a small chance that the rate hike will occur) or expectation that the T-bill rate will rise for some other reason, to be conservative I require her quarterly forecast $i^{\tau}_{j,t}$ to be more than 16 basis points above $i^{t-1}_{j,t}$, *and* that her forecast for the subsequent period $(i^{\tau}_{j,t})$ be at least 25 basis points above $i^{t-1}_{j,t}$, to determine that she expects liftoff in quarter τ. If I do not find that a forecaster in quarter t expects liftoff in quarters t, $t+1, \ldots, t+4$, then I record expected liftoff in year y if the corresponding annual forecast is more than 10 basis points above the backcast and the forecast for the subsequent calendar year is at least 25 basis points above the backcast.[30]

Figure A.1 shows that in 2010 and prior to the calendar-based guidance in 2011, the median SPF forecaster expects liftoff in two to three quarters. Note that the Blue Chip forecasts are for the quarterly average of the federal funds rate over the quarter indicated. Since Swanson and Williams require this forecast to be above 25 basis points to determine that a forecaster expects liftoff, they

[30]The modal difference between $i^{t+1}_{j,t}$ and $i^{t-1}_{j,t}$ in 2012 and 2013 is 0, so most of the time, if forecasters expect no change in the policy rate, they also forecast no change in the T-bill rate. Thus results are not very sensitive to the exact cutoff choices (figure A.2). Results are also unchanged if, instead of the backcast, I use the actual T-bill rate in the current or previous quarter, or a fixed rate of 10 basis points, the approximate T-bill rate at the ZLB. If expected liftoff is in the last reported horizon, I do not require that the forecast for the subsequent year (which is not provided) be at least 25 basis points above the backcast. If the annual forecast is at least 25 basis points above the backcast, I record that her expected liftoff date is in quarter 2 of the indicated calendar year, otherwise quarter 4 of the indicated calendar year.

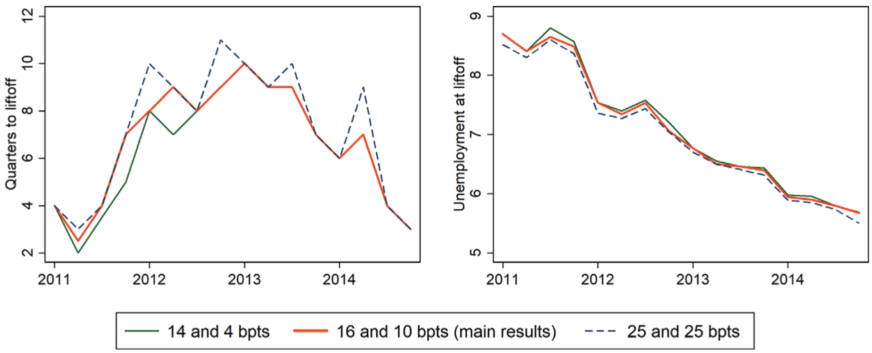
Figure A.1. Quarters to Expected Liftoff Date



Notes: Data are from SPF. The sample includes forecasters who provide expectations of three-month T-bill rate at all quarterly and annual horizons. Vertical lines indicate the starts of calendar-based and threshold-based forward guidance.

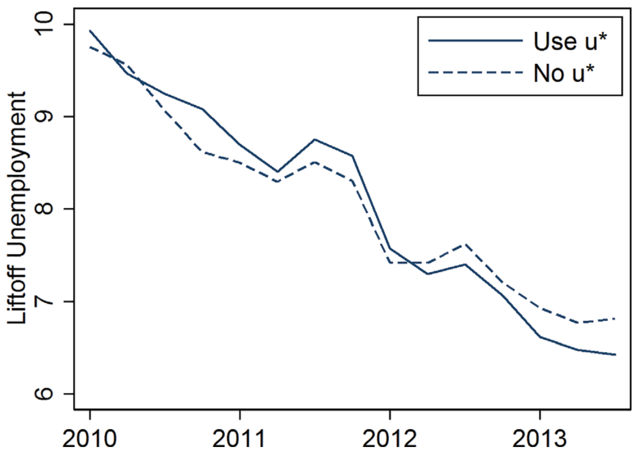
may slightly overestimate the time to expected liftoff. As discussed above, a 25 basis point rate hike near the middle or end of the quarter will be reflected in a quarterly average forecast of less than 25 basis points above the ZLB rate. Indeed, in figure A.2, which shows estimates of expected quarters to liftoff under different cutoffs, my results are more similar to those of Swanson and Williams (2014) when I use a 25 rather than 16 basis points cutoff to infer that a forecaster expects liftoff in a particular quarter. In the entire ZLB period, there are only three observations in which a forecaster expects liftoff beyond the longest forecast horizon. Moreover, there are only nine observations corresponding to expected liftoff in 2017 and two in 2018.

Figure A.2. Estimates of Expected Liftoff Timing and Unemployment under Alternative Cutoffs



Notes: Data are from SPF. The thick solid lines are the estimates used in this paper. See the text of appendix B for computational details.

Figure A.3. Estimates of Expected Liftoff Unemployment by Reported Use of u*



Notes: Data are from SPF. The solid line shows mean expected unemployment at time of first rate hike by forecasters who report using the natural rate of unemployment to forecast. The dashed line shows mean expected unemployment at time of first rate hike by forecasters who report not using the natural rate of unemployment to forecast.

Figure A.3 shows expected unemployment at liftoff over time for the mean forecaster who reports using or not using the natural rate of unemployment to forecast. The u^* users had slightly higher average liftoff unemployment expectations until 2012. By the end of 2012, when the FOMC introduced threshold-based forward guidance, the average u^* user's liftoff unemployment expectations were lower than that of the average non-user, and were quicker to drop below the 6.5 percent threshold.

References

Abraham, K. 2015. "Is Skill Mismatch Impeding U.S. Economic Recovery?" *ILR Review* 68 (2): 291–313.

Abraham, K., J. Haltiwanger, K. Sandusky, and J. Spletzer. 2019. "The Consequences of Long-Term Unemployment: Evidence from Linked Survey and Administrative Data." *ILR Review* 72 (2): 266–99.

Altig, D. 2012. "The Skills Gap: Still Trying to Separate Myth from Fact." Federal Reserve Bank of Atlanta Macroblog, June 7.

Andrade, P., R. Crump, S. Eusepi, and E. Moench. 2016. "Fundamental Disagreement." *Journal of Monetary Economics* 83 (October): 106–28.

Andrade, P., and H. Le Bihan. 2013. "Inattentive Professional Forecasters." *Journal of Monetary Economics* 60 (8): 976–82.

Andrade, P., G. Gaballo, E. Mengus, and B. Mojon. 2018. "Forward Guidance and Heterogeneous Beliefs." BIS Working Paper No. 750.

Ashley, R., K. P. Tsang, and R. J. Verbrugge. 2018. "All Fluctuations Are Not Created Equal: The Differential Roles of Transitory versus Persistent Changes in Driving Historical Monetary Policy." Working Paper No. 18-14, Federal Reserve Bank of Cleveland.

Ball, L. 2009. *Hysteresis in Unemployment: Old and New Evidence.* Federal Reserve Bank of Boston and MIT Press.

Bauer, M. D., and G. D. Rudebusch. 2013. "Expectations for Monetary Policy Liftoff." FRBSF Economic Letter No. 2013-34, Federal Reserve Bank of San Francisco.

Beqiraj, E., G. Di Bartolomeo, and M. Di Pietro. 2019. "Beliefs Formation and the Puzzle of Forward Guidance Power." *Journal of Macroeconomics* 60 (June): 20–32.

Bernanke, B. 2003. “A Perspective on Inflation Targeting.” Remarks at the Annual Washington Policy Conference of the National Association of Business Economists, Washington, DC, March 25.

———. 2007. “Inflation Expectations and Inflation Forecasting.” Remarks at the Monetary Economics Workshop of the National Bureau of Economic Research Summer Institute, Cambridge, Massachusetts, July 10.

———. 2016a. “Federal Reserve Economic Projections: What Are They Good For?” Brookings Blog, November 28.

———. 2016b. “The Fed’s Shifting Perspective on the Economy and Its Implications for Monetary Policy.” Brookings Blog, August 8.

Binder, C. 2017a. “Fed Speak on Main Street: Central Bank Communication and Household Expectations.” *Journal of Macroeconomics* 52 (June): 238–51.

———. 2017b. “Federal Reserve Communication and the Media.” *Journal of Media Economics* 30 (4): 191–214.

———. 2017c. “Measuring Uncertainty Based on Rounding: New Method and Application to Inflation Expectations.” *Journal of Monetary Economics* 90 (October): 1–12.

Binder, C., and S. Wetzel. 2018. “The FOMC versus the Staff, Revisited: When Do Policymakers Add Value?” *Economics Letters* 171 (October): 72–75.

Blanchard, O. 2018. “Should We Reject the Natural Rate Hypothesis?” *Journal of Economic Perspectives* 32 (1): 97–120.

Blinder, A. 2010. “Quantitative Easing: Entry and Exit Strategies.” *Review* (Federal Reserve Bank St. Louis) 92 (6): 465–79.

———. 2018. “Through a Crystal Ball Darkly: The Future of Monetary Policy Communication.” *American Economic Review, Papers and Proceedings* 108 (May): 567–71.

Bobeica, E., and M. Jarocinski. 2019. “Missing Disinflation and Missing Inflation: A VAR Perspective.” *International Journal of Central Banking* 15 (1, March): 199–232.

Boero, G., J. Smith, and K. Wallis. 2008. “Uncertainty and Disagreement in Economic Prediction: The Bank of England Survey of External Forecasters.” *Economic Journal* 118 (530): 1107–27.

Boneva, L., R. Harrison, and M. Waldron. 2018. “Threshold-based Forward Guidance.” *Journal of Economic Dynamics and Control* 90 (May): 138–55.

Borio, C., P. Disyatat, and M. Juselius. 2017. "Rethinking Potential Output: Embedding Information about the Financial Cycle." *Oxford Economic Papers* 69 (3): 655–77.

Branch, W., J. Carlson, G. Evans, and B. McGouch. 2004. "Monetary Policy, Endogenous Inattention, and the Volatility Tradeoff." Working Paper No. 04-11, Federal Reserve Bank of Cleveland.

Brock, W. A., and C. H. Hommes. 1997. "A Rational Route to Randomness." *Econometrica* 65 (5): 1059–95.

Bundick, B., and T. Herriford. 2017. "How Do FOMC Projections Affect Policy Uncertainty?" *Economic Review* (Federal Reserve Bank of Kansas City) (QII): 5–22.

Busetti, F., D. Delle Monache, A. Gerali, and A. Locarno. 2017. "Trust, but Verify: De-anchoring of Inflation Expectations under Learning and Heterogeneity." ECB Working Paper No. 1994.

Campbell, J. R., F. Ferroni, J. D. M. Fisher, and L. Melosi. 2019. "The Limits of Forward Guidance." *Economic Perspectives* (Federal Reserve Bank of Chicago) 43 (2): 2–12.

Campbell, J. R., J. D. M. Fisher, C. L. Evans, and A. Justiniano. 2012. "Macroeconomic Effects of Federal Reserve Forward Guidance." *Brookings Papers on Economic Activity* (Spring): 1–80.

Clarida, R. 2019. "Sustaining Maximum Employment and Price Stability." Speech at the Economic Club of New York, May 30.

Coibion, O., and Y. Gorodnichenko. 2012. "What Can Survey Forecasts Tell Us about Information Rigidities?" *Journal of Political Economy* 120 (1): 116–59.

———. 2015. "Is the Phillips Curve Alive and Well after All? Inflation Expectations and the Missing Disinflation." *American Economic Journal: Macroeconomics* 7 (1): 197–232.

Cole, S. J. 2015. "Learning and the Effectiveness of Central Bank Forward Guidance." Working Paper, Marquette University.

Cornea-Madeira, A., C. Hommes, and D. Massaro. 2019. "Behavioral Heterogeneity in U.S. Inflation Dynamics." *Journal of Business and Economic Statistics* 37 (2): 288–300.

Davis, J. S. 2012. "Inflation Expectations Have Become More Anchored Over Time." *Economic Letters* (Federal Reserve Bank of Dallas) 7 (13): 1–4.

Del Negro, M., M. Giannoni, and C. Patterson. 2013. "The Forward Guidance Puzzle." Staff Report No. 574, Federal Reserve Bank of New York.

DeLong, B. 1997. *America's Only Peacetime Inflation: The 1970s.* University of Chicago Press.

Detmers, G.-A. 2016. "Forward Guidance under Disagreement: Evidence from the Fed's Dot Projections." SFB 649 Discussion Paper No. 041.

Drager, L., M. J. Lamla, and D. Pfajfar. 2016. "Are Survey Expectations Theory-Consistent? The Role of Central Bank Communication and News." *European Economic Review* 85 (June): 84–111.

Eggertsson, G. 2006. "The Deflation Bias and Committing to Being Irresponsible." *Journal of Money, Credit and Banking* 38 (2): 283–322.

Eggertsson, G., and M. Woodford. 2003. "The Zero Bound on Interest Rates and Optimal Monetary Policy." *Brookings Papers on Economic Activity* 34 (1): 139–235.

Elsby, M. W., B. Hobijn, and A. Sahin. 2010. "The Labor Market in the Great Recession." *Brookings Papers on Economic Activity* 41 (Spring): 1–69.

Engen, E. M., T. Laubach, and D. Reifschneider. 2015. "The Macroeconomic Effects of the Federal Reserve's Unconventional Monetary Policies." Finance and Economics Discussion Series No. 2015-005, Board of Governors of the Federal Reserve System.

Espinosa-Vega, M. A., and S. Russell. 1997. "History and Theory of the NAIRU: A Critical Review." *Economic Review* (Federal Reserve Bank of Atlanta) 82 (Q2): 4–25.

Estrella, A., and F. S. Mishkin. 1999. "Rethinking the Role of NAIRU in Monetary Policy: Implications of Model Formulation and Uncertainty." In *Monetary Policy Rules*, ed. J. B. Taylor, 405–30. University of Chicago Press.

Eusepi, S., and B. Preston. 2018. "The Science of Monetary Policy: An Imperfect Knowledge Perspective." *Journal of Economic Literature* 56 (1): 3–59.

Faberman, R. J., and B. Mazumder. 2012. "Is There a Skills Mismatch in the Labor Market?" Chicago Fed Letter No. 300 (July), Federal Reserve Bank of Chicago.

Falck, E., M. Hoffmann, and P. Hurtgen. 2017. "Disagreement and Monetary Policy." Discussion Paper No. 29/2017, Bundesbank.

Farmer, R. E. A. 2013. "The Natural Rate Hypothesis: An Idea Past Its Sell-by Date." NBER Research Working Paper No. 19267.

Faust, J. 2016. "Oh What a Tangled Web We Weave: Monetary Policy in Divisive Times." Speech at the "Understanding Fedspeak" event co-sponsored by the Hutchins Center on Fiscal and Monetary Policy at the Brookings Institution and the Center for Financial Economics at Johns Hopkins University, Washington, DC, November 30.

Ferrero, G., and A. Secchi. 2010. "Central Banks' Macroeconomic Projections and Learning." Working Paper No. 782, Banca d'Italia.

Friedman, M. 1968. "The Role of Monetary Policy." *American Economic Review* 58 (1): 1–17.

Goy, G., C. Hommes, and K. Mavromatis. 2018. "Forward Guidance and the Role of Central Bank Credibility under Heterogeneous Beliefs."

Hayo, B., and M. Neuenkirch. 2013. "Do Federal Reserve Presidents Communicate with a Regional Bias?" *Journal of Macroeconomics* 35 (March): 62–72.

Holston, K., T. Laubach, and J. C. Williams. 2017. "Measuring the Natural Rate of Interest: International Trends and Determinants." *Journal of International Economics* 108 (Supplement 1): S59–S75.

Hommes, C., and J. Lustenhouwer. 2019. "Inflation Targeting and Liquidity Traps under Endogenous Credibility." *Journal of Monetary Economics* 107 (November): 48–62.

Honkapohja, S., and K. Mitra. 2015. "Comparing Inflation and Price-Level Targeting: The Role of Forward Guidance and Transparency." *The Manchester School* 83 (S2): 27–59.

Jung, A., and S. Latsos. 2015. "Do Federal Reserve Bank Presidents Have a Regional Bias?" *European Journal of Political Economy* 40 (Part A): 173–83.

Krugman, P. 1998. "It's Baaack: Japan's Slump and the Return of the Liquidity Trap." *Brookings Papers on Economic Activity* 2: 137–206.

Kumar, S., H. Afrouzi, O. Coibion, and Y. Gorodnichenko. 2015. "Inflation Targeting Does Not Anchor Inflation Expectations: Evidence from Firms in New Zealand." *Brookings Papers on Economic Activity* (Fall): 151–225.

Laubach, T., and J. C. Williams. 2016. “Measuring the Natural Rate of Interest Redux.” *Business Economics* 51 (2): 57–67.

Mankiw, N. G., R. Reis, and J. Wolfers. 2004. “Disagreement about Inflation Expectations.” In *NBER Macroeconomics Annual 2003*, Vol. 18, ed. M. Gertler and K. Rogoff, 209–48. MIT Press.

McKay, A., E. Nakamura, and J. Steinsson. 2016. “The Power of Forward Guidance Revisited.” *American Economic Review* 106 (10): 3133–58.

Meade, E., and D. Thornton. 2012. “The Phillips Curve and US Monetary Policy: What the FOMC Transcripts Tell Us.” *Oxford Economic Papers* 64 (2): 197–216.

Morris, S., and H.-S. Shin. 2002. “The Social Value of Public Information.” *American Economic Review* 92 (5): 1521–34.

Nechio, F., and R. Regan. 2016. “Fed Communication: Words and Numbers.” FRBSF Economic Letter No. 2016-26, Federal Reserve Bank of San Francisco.

Olson, P., and D. Wessel. 2016. “Federal Reserve Communications: Survey Results.” Report from Hutchins Center on Fiscal and Monetary Policy at Brookings.

Orphanides, A. 2001. “Monetary Policy Rules Based on Real-Time Data.” *American Economic Review* 91 (4): 964–85.

———. 2019. “Monetary Policy Strategy and Its Communication.” Paper presented at the Federal Reserve Bank of Kansas City Economic Policy Symposium, Jackson Hole, Wyoming, August 22–24.

Orphanides, A., and J. Williams. 2002. “Robust Monetary Policy Rules with Unknown Natural Rates.” *Brookings Papers on Economic Activity* 2: 63–145.

———. 2005. *Imperfect Knowledge, Inflation Expectations, and Monetary Policy.* Chicago: University of Chicago Press.

———. 2006. “Monetary Policy with Imperfect Knowledge.” *Journal of the European Economic Association* 4 (2–3): 366–75.

Patton, A., and A. Timmermann. 2010. “Why Do Forecasters Disagree? Lessons from the Term Structure of Cross-sectional Dispersion.” *Journal of Monetary Economics* 57 (7): 803–20.

Pierdzioch, C., J.-C. Rulke, and G. Stadtmann. 2011. “Do Professional Economists’ Forecasts Reflect Okun’s Law? Some Evidence for the G7 Countries.” *Applied Economics* 43 (11): 1365–73.

Powell, J. 2016. “A View from the Fed.” Speech at the “Understanding Fedspeak” event co-sponsored by the Hutchins Center on Fiscal and Monetary Policy at the Brookings Institution and the Center for Financial Economics at Johns Hopkins University, Washington, DC, November 30.

Primiceri, G. E. 2006. “Why Inflation Rose and Fell: Policy-Makers’ Beliefs and U. S. Postwar Stabilization Policy.” *Quarterly Journal of Economics* 121 (3): 867–901.

Reis, R. 2003. “Where Is the Natural Rate? Rational Policy Mistakes and Persistent Deviations of Inflation from Target.” *Advances in Macroeconomics* 3 (1): 1–38.

Romer, C., and D. Romer. 2002. *The Evolution of Economic Understanding and Postwar Stabilization Policy.* Federal Reserve Bank of Kansas City.

———. 2004. “Choosing the Federal Reserve Chair: Lessons from History.” *Journal of Economic Perspectives* 18 (1): 129–62.

———. 2008. “The FOMC versus the Staff: Where Can Monetary Policymakers Add Value?” *American Economic Review* 98 (2): 230–35.

Rudebusch, G. D., and L. E. O. Svensson. 1999. “Policy Rules for Inflation Targeting.” In *Monetary Policy Rules*, ed. J. B. Taylor, 203–53 (chapter 5). Chicago: University of Chicago Press.

Rulke, J.-C. 2012. “Do Professional Forecasters Apply the Phillips Curve and Okun’s Law? Evidence from Six Asian-Pacific Countries.” *Japan and the World Economy* 24 (4): 317–24.

Sargent, T., N. Williams, and T. Zha. 2006. “Shocks and Government Beliefs: The Rise and Fall of American Inflation.” *American Economic Review* 96 (4): 1193–1224.

Simon, H. A. 1984. “On the Behavioral and Rational Foundations of Economic Dynamics.” *Journal of Economic Behavior and Organization* 5 (1): 35–55.

Staiger, D., J. H. Stock, and M. W. Watson. 1997. “The NAIRU, Unemployment and Monetary Policy.” *Journal of Economic Perspectives* 11 (1): 33–49.

Stasavage, D. 2007. “Polarization and Publicity: Rethinking the Benefits of Deliberative Democracy.” *Journal of Politics* 69 (1): 59–72.

Sunstein, C. R. 2017. “Output Transparency vs. Input Transparency.” Manuscript.

Swanson, E. T., and J. C. Williams. 2014. “Measuring the Effect of the Zero Lower Bound on Medium- and Longer-Term Interest Rates.” *American Economic Review* 104 (10): 3154–85.

Tasci, M., and R. J. Verbrugge. 2014. “How Much Slack Is in the Labor Market? That Depends on What You Mean by Slack.” Economic Commentary No. 2014-21 (October 15), Federal Reserve Bank of Cleveland.

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

Thornton, D. L. 2003. “Monetary Policy Transparency: Transparent About What?” *The Manchester School* 71 (5): 478–97.

———. 2012. “How Did We Get to Inflation Targeting and Where Do We Need to Go to Now? A Perspective from the U.S. Experience.” *Review* (Federal Reserve Bank of St. Louis) 94 (Jan/Feb): 65–81.

———. 2015. “Requiem for QE.” Policy Analysis No. 783, Cato Institute.

Walsh, C. E. 1998. “The Natural Rate, NAIRU, and Monetary Policy.” FRBSF Economic Letter No. No. 1998-28, Federal Reserve Bank of San Francisco.

Williams, J. C. 2006. “Robust Estimation and Monetary Policy with Unobserved Structural Change.” *FRBSF Economic Review* (2006): 1–16.

———. 2013a. “Forward Policy Guidance at the Federal Reserve.” VoxEU.

———. 2013b. “Will Unconventional Policy Be the New Normal?” FRBSF Economic Letter No. 2013-29, Federal Reserve Bank of San Francisco.

Woodford, M. 2012. “Methods of Policy Accommodation at the Interest-Rate Lower Bound.” Federal Reserve Bank of Kansas City Symposium on The Changing Policy Landscape, Jackson Hole, Wyoming, August 30–September 1.

Yagan, D. 2019. “Employment Hysteresis from the Great Recession.” *Journal of Political Economy* 127 (5): 2505–58.

Yellen, J. 2012. “Revolution and Evolution in Central Bank Communications.” Speech at the Haas School of Business, University of California, Berkeley, November 13.

What Does “Below, but Close to, 2 Percent” Mean? Assessing the ECB’s Reaction Function with Real-Time Data*

Maritta Paloviita, Markus Haavio, Pirkka Jalasjoki, and Juha Kilponen
Bank of Finland

Using unique real-time quarterly macroeconomic projections of the Eurosystem/ECB staff, we estimate competing specifications of the ECB’s monetary policy reaction function. We consider specifications which include inflation and output growth projections, a past inflation gap, a time-varying natural real interest rate, and different inflation targets. Our first key finding is that the de facto inflation target of the ECB lies between 1.6 percent and 1.8 percent. Our second key finding is that the ECB reacts both to short-term macroeconomic projections and to past deviations of inflation from its de facto target.

JEL Codes: E31, E52, E58.

*All authors are from the Bank of Finland. We are indebted to Giacomo Carboni, Seppo Honkapohja, Aki Kangasharju, Jarmo Kontulainen, Tomi Kortela, Mika Kortelainen, Athanasios Orphanides, Antonio Rua, and Tuomas Välimäki for valuable comments and suggestions. We are also grateful for constructive suggestions received at a seminar organized by the National Bank of Belgium, the XL Annual Meeting of the Finnish Economic Association, ECB Working Group of Forecasting meeting, Bank of Russia Conference on Inflation: New Insights for Central Banks, 15th Euroframe Conference on Economic Policy Issues in the European Union, Cracow University of Economics Workshop on Macroeconomic Research, MMF 50th Annual Conference, ECB Monetary Policy Committee Meeting, National Bank of Poland Research Seminar, Bank of Finland Monetary Policy Seminar, and Nordic Monetary Policy Meeting. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Bank of Finland or the Eurosystem. Corresponding author e-mail: maritta.paloviita@bof.fi.

1. Introduction

In recent years, inflation has been persistently low in many economies. As a response, policy rates have been cut to very low levels, and new measures have been introduced to maintain an accommodative stance of monetary policy. The low inflation and interest rate environment have raised the question of whether and how the current monetary policy framework should be reformed (see, e.g., Bernanke 2017a, 2017b; Williams 2017; Bullard 2018; Honkapohja and Mitra 2018). In the case of the European Central Bank, there has also been a vivid debate on the precise numerical target for inflation and possible asymmetry of the ECB's policy responses.

The debate on the ECB's price stability objective stems from the fact that its inflation aim is not precisely defined in the Treaty on the Functioning of the European Union. In 1998, the ECB's Governing Council defined price stability as a "year-on-year increase in the Harmonised Index of Consumer Prices (HICP) for the euro area of below 2%." In 2003, the Governing Council clarified that "in the pursuit of price stability it aims to maintain inflation rates below, but close to, 2% over the medium term." This clarification can be seen as an effort to reduce uncertainty about the lower bound of the inflation aim relative to the earlier definition and to provide a buffer against large negative shocks to inflation.

As discussed in Hartmann and Smets (2018), the exact formulation by the Governing Council in 2003 was a compromise that maximized that buffer while remaining consistent with the definition of price stability. With this reformulation, the inflation aim remained nevertheless ambiguous.[1] In particular, although the ECB communication stresses symmetry, the expression "below, but close to 2%" has some feel of asymmetry, and the exact numerical target is not spelled out.[2]

[1] Apel and Claussen (2017) classify three different categories for inflation targeting. A point target refers to a single number. If certain deviations from the point target are "acceptable" for the central bank, it is complemented with a tolerance band. In the case of a target range, a targeted inflation interval is announced without any specific desirable level of inflation.

[2] According to the ECB strategy, "the Governing Council's aim to keep euro area inflation below, but close to, 2% over the medium term signifies a commitment to avoiding both inflation that is persistently too high and inflation

Not surprisingly, the ECB's inflation aim has been interpreted in various ways. For example, Miles et al. (2017) point out that the ECB's "target itself is perceived as asymmetric." They also note that there "is uncertainty about what 'close to, but below' means." Regarding the public, survey evidence indicates that households' knowledge about the ECB's inflation target is "far from perfect" (van der Cruijsen, Jansen, and de Haan 2015). Different interpretations of the inflation target and/or vague monetary policy communication may increase inefficiency in monetary policymaking, give rise to risks of deanchored inflation expectations, and, hence, jeopardize the effective transmission of monetary policy. After introducing new policy measures, the ECB has strengthened its communication and adopted a forward-guidance framework in order to reduce uncertainty concerning its reaction function and future policy actions.[3]

In this paper, we are specifically interested in assessing the ECB's own interpretation of the price stability objective and its reaction function. Using unique real-time quarterly macroeconomic projections of the Eurosystem/ECB staff, we attempt to quantify the gist of the expression "below, but close to 2%." First, we consider the levels toward which the ECB inflation projections converge in the medium term. Second, we estimate a large number of alternative output-growth-gap-based reaction functions in order to directly infer the ECB's de facto inflation target. Finally, using primarily the real gross domestic product (GDP) growth as a cyclical variable, we estimate more general reaction function specifications, which allow the ECB to react (either symmetrically or asymmetrically) also to past inflation gaps, determined by the deviations of realized inflation from the de facto target. In all cases, we pay special attention to the relevant forecast horizon in monetary policymaking.

A novel feature of our analysis is that our data set includes the Eurosystem/ECB staff quarterly macroeconomic projections of inflation and real GDP growth made in 1999–2016. Consequently,

---

that is persistently too low." For example, in March 2016 Mario Draghi, President of the European Central Bank, stated: "The key point is that the Governing Council is symmetric in the definition of the objective of price stability over the medium term." (https://www.ecb.europa.eu/press/pressconf/2016/html/is160310.en.html.) See also President Draghi's speech of June 2016, at https://www.ecb.europa.eu/press/key/data/2016/html/sp160602.en.html.

[3]See, e.g., Cœuré (2017).

we are able to estimate the reaction function with a subset of the very same information the Governing Council has available when it decides on the monetary policy stance.[4] As emphasized by Woodford (2007), an important feature of "optimal" monetary policy is that it should respond to the projected future path of the economy and not only to current conditions.

Our sample period, 1999:Q4–2016:Q4, covers the relatively stable pre-crisis years as well as the recent turbulent years characterized by the financial crisis, the sovereign debt crisis, and low inflation. Using subsample analysis and recursive estimations, we analyze the stability of estimated parameters of the ECB's reaction function over time. In addition to the targeted rate of inflation, we conduct a robustness analysis with respect to the time span of forward-looking and backward-looking variables in the reaction function and with respect to time-varying long-run natural real interest rates. We assess the performance of estimated reaction functions by comparing their in-sample predictions against the key interest rates. In the analysis of the most recent period when standard interest rate policy has approached its effective lower bound, we evaluate the performance of our estimated functions by comparing their out-of-sample predictions against shadow interest rates estimated by Kortela (2016) and by Wu and Xia (2016).

Our extensive analyses based on alternative approaches and unique real-time data indicate that the de facto inflation target of the Governing Council lies between 1.6 percent and 1.8 percent. This finding is consistent with the fact that the Eurosystem/ECB staff medium-term inflation projections have had a tendency to converge rapidly on values well below 2 percent. We also find that the ECB conditions its interest rate decisions not only on short-term macroeconomic projections but also on past inflation developments. This is also consistent with the recent ECB communication, according to which the launch of asset purchase programs can be justified as a response to too-prolonged a period of low inflation. Finally, we find

[4]To our knowledge, earlier reaction function estimations using the ECB's projections have been based on annual information only, with one exception: Hartmann and Smets (2018). Fischer et al. (2009) examine euro-area monetary analysis in 1999–2006 using quarterly information.

some evidence of asymmetry in policy rules in which we fix the inflation target to 2 percent. However, the out-of-sample predictions of the symmetric reaction function with a low de facto target outperform the asymmetric reaction function during the zero lower bound period.

In earlier studies, euro-area monetary policy has also been widely examined using alternative specifications of the classical Taylor rule (Taylor 1993).[5] Monetary policy analysis has often been based on real-time information. As a proxy for real-time information, the ECB Survey of Professional Forecasters (ECB SPF) (e.g., Gerlach and Lewis 2014) and Consensus Forecast (e.g., Gorter, Jacobs, and de Haan 2008) have been used. Some authors have also used the ECB's macroeconomic projections (e.g., Belke and Klose 2011; Bletzinger and Wieland 2017). As the ECB projections were published only for full calendar years until 2017, quarterly variation in the projections has been taken into account in reaction function estimations so far only by Hartmann and Smets (2018). Close to our study also is an article by Bletzinger and Wieland (2017), who also estimate a forecast-based reaction function for the euro area in order to assess the targeted level of inflation and the ECB policy during the zero lower bound period. Their analysis is based on the ECB SPF survey and the ECB projections for full calendar years. The main difference from our approach is that they do not take into account the impact of past inflation deviations from the target, and their cyclical variable is defined as a difference between output growth and the European Commission's estimate of potential output growth.[6]

The paper is organized as follows. The Eurosystem/ECB staff projections are described and their medium-term convergence is examined in section 2. Alternative specifications of the monetary policy reaction function and estimation results are presented in section 3. In section 4, we discuss in-sample and out-of-sample predictions of different reaction functions. Concluding remarks are provided in section 5.

[5]See, e.g., Clarida, Galí, and Gertler (1998).

[6]Earlier studies of possible asymmetries in ECB monetary policy include Surico (2003, 2007), Aguiar and Martins (2008), and Ikeda (2010).

2. The Data and Eurosystem/ECB Staff Projections

2.1 Data Description

Our data set includes the real-time Eurosystem/ECB staff projections made in 1999:Q4–2016:Q4 for the euro-area year-on-year HICP inflation rate and year-on-year real GDP growth rate. These projections are publicly available as annual data for full calendar years, but our analyses are based on confidential quarterly information.[7]

For both the inflation rate and real GDP growth rate our data include real-time estimates of previous-quarter values, current-quarter values (nowcast estimates), and real-time projections until the end of each forecast horizon. The projections in our data cover the current and next two calendar years. The “final” data, i.e., revised data, for our purposes, are the latest available vintages published by Eurostat in the spring of 2017. The euro-area GDP data are seasonally and working-day adjusted.

Projection errors increase substantially with the length of the forecast horizon, reflecting real-time challenges in actual monetary policymaking.[8] The mean errors (ME) for the one- to four-quarters-ahead inflation projections (real GDP growth projections) are −0.02, −0.06, −0.11, and −0.13 (−0.11, 0.07, 0.29, and 0.51). The corresponding root mean squared errors (RMSE) for inflation are 0.37, 0.59, 0.78, and 0.95 and for real GDP growth 0.96, 1.33, 1.68, and 1.96. A limited forecast accuracy in the medium to long term is not specific to the ECB and the Eurosystem. Charemza and Ladley (2016) have analyzed inflation forecasts made in 2000–11 in 10 inflation-targeting central banks (the ECB is not included in the study). They show that compared with the CESifo World Economic Survey forecasts, the central banks’ one-year-ahead inflation forecasts are biased toward the inflation target. According to their analysis, the bias is even stronger in two-years-ahead inflation forecasts.[9]

---

[7]See Alessi et al. (2014) and ECB (2016) for a detailed description of the Eurosystem/ECB staff projections exercises.

[8]We define projection errors as the difference between projections and realizations.

[9]Charemza and Ladley’s (2016) analysis includes 10 inflation-targeting central banks: Australia, Canada, Chile, Czech Republic, Korea, Mexico, New Zealand,

Our data set also includes the EONIA (euro overnight index average) interest rate, the MRO (main refinancing operations) rate, and shadow interest rates estimated by Kortela (2016) and Wu and Xia (2016).[10] The shadow rates follow closely the EONIA rate until about mid-2014, but thereafter they start to fall strongly into a negative territory, reflecting the quantitative easing of the ECB (see figure A.1 in appendix A).

We also calculate several time-varying ex ante and ex post proxies of the long-run natural real interest rate, which are constructed using yields on German government bonds of different maturities or a composite nominal yield of 10-year euro-area government bonds. The composite nominal yield is constructed by the ECB by aggregation using GDP weights. We use these different proxies of the natural rate because of measurement issues. Differences in long-term bond yields of different euro-area economies were small until around the inception of global financial crisis in our sample, so it does not make a great difference whether the German government bonds or composite yield is used for that period. Since about 2007, however, the difference becomes significant, and the German government bond yields are likely to be a better proxy for the euro-area risk-free nominal rate, as it corresponds to the lowest of the 10-year government bond yields. However, this is not necessarily the best proxy for the euro-area long-run natural rate, i.e., for the rate which would stabilize the euro-area economy as a whole, in the long run. The literature in general uses either short-term bond yields or long-term bond yields to approximate the natural rate.

Norway, Poland, and Sweden. See Sveriges Riksbank (2017) for the accuracy of Riksbank's inflation forecasts.

[10] A shadow rate is a summary measure of monetary policy stance, capturing unconventional as well as conventional policy measures. It indicates how much the central bank would have lowered the interest rates had the zero lower bound not been binding, i.e., it reflects monetary policy stance in very low or negative interest rate environments. Differences between alternative shadow rates for the euro area based on different methods are typically quite large. However, they all indicate that the ECB's monetary policy stance has recently been very accommodative. In Kortela (2016), the shadow rate is based on a multifactor shadow rate term structure model (SRTSM) with a time-varying lower bound.

2.2 Medium-Term Convergence of Inflation and Output Growth Projections

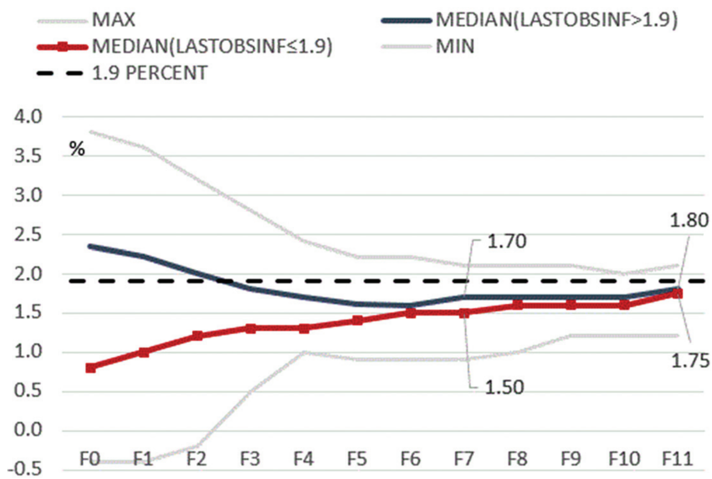
The Eurosystem/ECB staff inflation projections provide suggestive evidence of the ECB's de facto inflation target. Projected values of the economic variables, including inflation, at the end of the forecast horizon are largely determined by the models' long-run equilibriums, i.e., values to which they are expected to converge in the absence of new shocks hitting the economy. This is important for the determination of inflation itself, since empirical literature largely agrees that the central bank forecasts have an impact on the private sector's inflation forecasts and expectations.[11] It is important to note that the ECB inflation projections are conditioned on market expectations of the interest rate (since June 2006) and not on some "optimal state contingent path of the interest rates." Therefore, the projected inflation does not reflect the ECB's desired path of inflation per se. However, one can plausibly argue that the projected inflation rates at the end of the forecast horizon give the public a good guideline for inflation which the ECB considers consistent with its mandate. This is supported by the fact that inflation forecasts have typically converged to the promixity of "close but below two" already after about six quarters. There is rather little movement in inflation forecasts thereafter, as we show below.

Figure 1 illustrates the inflation projections. It shows two separate medians of the inflation projections based on whether the latest observed inflation rate during each projection exercise has been above or below 1.9 percent. More precisely, we have organized the projection data in figure 1 in the following way: the label "F0" on the horizontal axis refers to the median value of nowcast estimates from all the projection vintages and the labels "F1"–"F11" refer to the median values of the corresponding inflation projections for 1–11 quarters ahead. In addition to the medians, figure 1 also presents the highest and lowest inflation projections for different forecast horizons.

Figure 1 shows that the medians of projections made at times when the recent observed inflation rate is high (i.e., higher than

[11]See, e.g., Fujiwara (2005), Hubert (2015), and Lyziak and Paloviita (2018).

Figure 1. Median Inflation Projections Conditioned on Latest Observed Inflation Rate during Each Projection Exercise

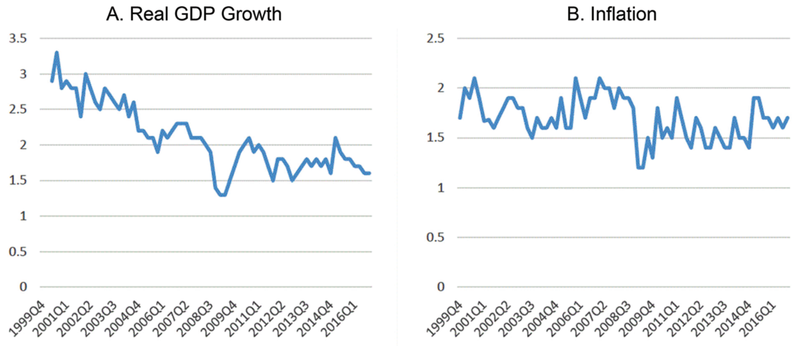


Sources: ECB and authors' own calculations.
Notes: On the horizontal axis, the label "F0" refers to the real-time current-quarter nowcasts and the label "F1" to the one-quarter-ahead projections, etc. The curves "MAX" and "MIN" refer to the highest and lowest inflation projections made in 1999:Q4–2016:Q4.

1.9 percent) converge to 1.70–1.80 percent after six quarters. At the same time, however, the medians of projections starting from lower inflation conditions (i.e., 1.9 percent or lower) converge to slightly lower rates around 1.60–1.75 percent. Lower medians converge to their eventual rates in a rather linear fashion, while the evolution of the higher medians has a somewhat different shape: the median projections for five and six quarters ahead are slightly below the medians at the end of the forecast horizon, i.e., inflation is projected to temporarily undershoot when inflation has been initially above 1.9 percent.

It is notable in figure 1 that regardless of the current level of inflation, after about six quarters the median inflation projections are already in the proximity of their levels at the end of the forecast horizon. When compared with the actual realized inflation, the

Figure 2. Medium-Term Projections



Sources: ECB and authors' own calculations.
Note: The graphs present the real GDP growth and inflation projections for the last quarter of the projection horizon for each projection exercise (the horizontal axis).

projected inflation exhibits stronger and faster mean reversion. Similarly to the inflation forecasts, the GDP growth forecasts also have a tendency to revert very quickly back to the perceived long-run growth rate. As a result, in both cases of inflation and GDP growth, the sample standard errors are much higher than the standard errors computed from different forecast vintages, especially at the end of the forecast horizon (see table A.1 in appendix A).

The medium-term real GDP growth and inflation projections are summarized in figure 2. The GDP growth projections do not revert toward a single long-run value over the sample, but rather the projections seem to capture the slowdown of long-run growth rates over the sample period. While at the beginning of the sample the GDP growth projections converged to growth rates of around 2.5 percent, more recently the projections have converged to below 2 percent growth. This decline in the projected medium-term growth rate is consistent with the trendlike decline in the real interest rates (see figure A.2 in appendix A), and also with the more recent Eurosystem's view that the potential growth of the euro-area economy is in the proximity of 1.5 percent. In contrast to the medium-term GDP growth forecasts, the inflation forecasts do not show a similar downward trend.

3. Estimation of the ECB Reaction Function

In what follows, we estimate alternative specifications of the Eurosystem/ECB's reaction functions for the period 1999:Q4–2014:Q2 (i.e., until the zero lower bound was reached) and assess the ECB's de facto inflation target both directly and indirectly. In our extensive analysis, we pay special attention to real-time data challenges and we also focus on possible backward-looking features of the monetary policy decisions.

We proceed in two steps. We first consider simple output-growth-gap-based (Taylor-type) reaction functions, which allow us to calculate the ECB's implied inflation target based on the estimated parameters. Then, in section 3.2, we consider less standard specifications where we use output growth as a cyclical variable due to the difficulty of estimating the output gap in real time and the fact that the ECB's communication is based more on the current and future output growth than on the output gap (see, e.g., Orphanides and van Norden 2002, Gerlach 2007, Orphanides 2008).[12] Using these specifications, we are able to assess the value of the ECB's de facto inflation target indirectly.[13]

When estimating nonstandard specifications of the reaction function in section 3.2, we consider possible backward-looking features in the ECB monetary policymaking by following Neuenkirch and Tillmann (2014)[14]: we augment our forward-looking specifications with a backward-looking inflation gap term, which measures how strongly actual inflation has deviated on average from the presumed inflation target in recent quarters. This past inflation gap—i.e., a "credibility loss term"—is specified as

$$CL_t = (\bar{\pi}_{t-1,t-q} - \pi^*)|\bar{\pi}_{t-1,t-q} - \pi^*|. \quad (1)$$

[12]For the euro area, the problem of reliable real-time output gap estimates is especially severe, due to a relatively short sample and methodological issues that arise from calculating the real-time output gap based on country aggregations (Marcellino and Musso 2011).

[13]Output-growth-based reaction functions have been analyzed by several authors. See, for example, Gorter, Jacobs, and de Haan (2008), Sturm and de Haan (2011), Gerlach and Lewis (2014), and Neuenkirch and Tillmann (2014).

[14]Neuenkirch and Tillmann (2014) analyze monetary policy in five inflation-targeting economies: Australia, Canada, New Zealand, Sweden, and the United Kingdom.

$\bar{\pi}_{t-1,t-q}$ refers to an average past inflation rate and q to the number of lags. The CL term is specified such that it penalizes both negative and positive deviations of average past inflation from the target symmetrically. For instance, if average past inflation is 1 percentage point below or above the target, in both cases the CL term gets the same absolute value but the sign is different. The absolute value term in equation (1) weights large deviations of inflation from the target more than small ones, hence the term is nonlinear. This nonlinear feature is needed to make indirect inference on the de facto inflation target.

This CL term, if found significant, introduces history dependence in the ECB policymaking. Past inflation developments may play a role in monetary policy setting because of various reasons. First, if the actual inflation rate has been below (above) the inflation target over a long period of time, the central bank may need to aim for a slightly faster (slower) rise in prices in the near future in order to achieve the inflation target in the medium term. This implies more accommodative (tighter) policy than what the current economic outlook would otherwise imply (see, e.g., Woodford 2007). Second, if inflation has persistently deviated from the target, the central bank may react more aggressively than would be required by information based on purely macroeconomic forecasts to maintain its credibility and commitment to the target. In this case, monetary policy aims to ensure that general (longer-term) inflation expectations remain anchored to the central bank's inflation target (see, e.g., Ehrmann 2015, Lyziak and Paloviita 2017). The third possible interpretation relates to unconventional monetary policy and, above all, to forward guidance: in the context of persistently low inflation, the central bank may promise to keep monetary policy accommodative even after monetary policy should be tightened according to the current economic outlook. This kind of forward guidance may appear in the reaction function as a link between the current policy rate and past inflation.[15]

[15] Cœuré (2017) argues that the ECB's forward guidance is based on a structural component that corresponds to the ECB reaction function and a variable component which consists of evolving economic outlook. According to him, the reaction function "includes the mapping of any desired monetary policy stance into instruments, such as policy rates and asset purchases."

Figure A.3 in appendix A presents the evolution of the ECB's CL term for the inflation targets of 1.7 percent and 2.0 percent, using seven lags over which the average past inflation $\bar{\pi}_{t-1,t-7}$ is measured. Both measures indicate that in the mid-2000s, past inflation gaps were minor, while more pronounced past inflation gaps are measured around 2002, 2009, 2011, and 2013, and again after 2014 when the nominal interest rate hit the lower bound and inflation slowed down persistently. The relatively large inflation gaps especially in the post-crisis period may have had a significant impact on the ECB's monetary policy.

Finally, at the end of this section, we evaluate the performance of estimated reaction functions by comparing their in-sample predictions against the key interest rates and their out-of-sample predictions against shadow interest rates estimated by Kortela (2016) and by Wu and Xia (2016).

All estimations are based on the generalized method of moments (GMM) with lags of regressors as instruments. We use the heteroskedasticity and autocorrelation corrected (HAC) (Newey and West 1987) GMM weighting matrix, which accounts for heteroskedasticity and serial autocorrelation in the estimated reaction function residuals. Use of the GMM in this context is motivated by the potential simultaneity of the right-hand-side variables of the reaction function. It is conceivable that the forecasts for inflation and the cyclical variable are affected by current monetary policy. In addition, our reaction function includes a proxy for the neutral rate of interest, which is measured subject to error. To the extent that these errors are correlated with other regressors, ordinary least squares (OLS) would give biased estimates.

3.1 Linear Reaction Functions

We start with estimating a large number of competing linear reaction functions, in which we use the real-time output growth gap as a proxy for the cyclical stance in the economy:

$$i_t = \rho i_{t-1} + (1-\rho)(\alpha + \beta_\pi \pi^f_{t+j|t} + \beta_y(\triangle y^f_{t+k|t} - \triangle y^*_t) + r^*_t). \quad (2)$$

While this reaction function is not an outcome of explicit optimization based on a structural model and the central bank's preferences, it is comparable to an inflation-forecast-targeting procedure, advocated by Svensson and Woodford (2005), as a way to implement optimal state-contingent policy.[16]

In equation (2), the MRO rate, the average EONIA rate, or the end-of-quarter EONIA rate i_t measures the monetary policy stance and the term i_{t-1} captures interest rate smoothing. The term $\pi^f_{t+j|t}$ refers to the ECB's projection of j-quarters-ahead HICP inflation and $\triangle y^f_{t+k|t}$ to the ECB's projection of k-quarters-ahead real GDP growth. Potential output growth ($\triangle y^*_t$) is proxied by long-run output growth projections. The underlying assumption is that the medium-run growth projection for the euro area corresponds to the assessed real time euro-area growth potential.[17]

In the original Taylor (1993) formulation, the neutral real interest rate is set to a constant, equal to 2 percent. This implies together with a 2 percent inflation target that the equilibrium nominal rate would be 4 percent. There is compelling evidence that equilibrium real interest rates are variable and have been trending downward both in the United States and in the euro area recently.[18] While the

[16]The Eurosystem/ECB staff projections were at first based on a constant interest rate assumption, but in order to further improve the quality and internal consistency of macroeconomic projections, both short-term and long-term interest rate assumptions have been based on market expectations since the June 2006 projection exercise. According to the ECB (2006), "this change is of a purely technical nature," which "does not imply any change in the ECB's monetary policy strategy or in the role of projections within it." We therefore interpret this change as if the internal forecasting procedure of the ECB had changed, but we don't expect a change in the reaction function itself.

[17]Another option would have been to use potential output estimates. However, the real-time estimates for euro-area potential output are only available from 2009:Q2 at a quarterly frequency and from 2006 at an annual frequency in the ECB projection data. It is also worth noting that in the ECB's New Area-Wide Model (NAWM), the reaction function has been specified in terms of deviations of output growth from its long-run empirical mean (Christoffel, Coenen, and Warne 2008).

[18]Our specification of the interest rate rule, which includes a proxy for the natural rate of interest, is akin also to Wicksell (1898), who argued that in order to maintain price stability, monetary policy should aim to track some measure of neutral rate determined purely by real factors (such as productivity of capital). King and Wolman (1999) and Woodford (2003) have shown that such a rule can result from optimizing central bank behavior in a standard New Keynesian model. In this formulation of the policy rule, when the equilibrium real rate rises, the central bank sets the interest rate higher so as to keep the output (growth) close to its equilibrium level (see also Cúrdia et al. 2015).

equilibrium real interest rate is difficult to estimate and is subject to large uncertainty, there is no reason why a time-varying equilibrium rate could not be incorporated into a policy rule.[19] In line with Clarida (2012) and Neuenkirch and Tillmann (2014), we append the reaction function with the long-term real interest rate as a proxy for the equilibrium real rate (r_t^*). We use yields on German government bonds of different maturities and calculate the real rate by subtracting either ex ante or ex post inflation from the nominal yield.[20]

In order to interpret the expression "below, but close to 2 percent," we need to solve the implicit inflation target in equation (2). Assuming that (expected) inflation is at its target level (π^*), output growth is at the potential level ($\triangle y_t^*$), the natural rate is constant (r^*), and the policy rate is constant over time (i), we can present the steady-state version of equation (2) in the following form: $i = \alpha + \beta_\pi \pi^* + r^*$. When combined with the Fisher equation ($i = \pi^* + r^*$), we can find the implicit inflation target $\pi^* = -\alpha/(\beta_\pi - 1)$.[21]

When estimating reaction functions, forecast horizons for forward-looking variables are typically assumed to be relatively short, reflecting a poorer forecast accuracy over a longer period of time.[22] However, we consider forward-lookingness of the ECB's policy responses without fixing forecasts horizons a priori by varying forecast horizons of inflation and output growth from zero (i.e., nowcast) to four quarters. Correspondingly, when constructing proxies for potential output growth, we use output growth projections from 8 to 11 quarters ahead.[23]

---

[19] For discussion see, e.g., Taylor (2018).

[20] Basing the proxy for the time-varying natural rate on German bunds instead of generic GDP-weighted composite euro-area bond yields is motivated in this context by the fact that German bund yields arguably do not contain the default risk premiums in the latter half of the sample.

[21] Since the implicit inflation target $\pi^* = -\alpha/(\beta_\pi - 1)$ is a ratio of estimation coefficients, its 95 percent confidence band can be computed with the help of the standard deviations of the coefficients α and β_π, and their correlation. Notice that the confidence band of the implicit inflation target is typically not symmetric.

[22] For example, when estimating reaction functions for the ECB, both Gerlach and Lewis (2014) and Neuenkirch and Tillmann (2014) consider one-year-ahead forecasts of inflation and output growth.

[23] We instrument our measure of potential output growth (eight-quarters-ahead growth forecast) with eight-quarters-ahead growth forecasts from past data

Among the resulting large number of reaction function candidates, we then use the following criteria to choose our preferred specifications in order to assess the ECB's de facto inflation target:

(i) The computed inflation target must have a bounded 95 percent confidence interval, which implies that the Taylor principle holds at the 5 percent level., i.e., $\beta_\pi > 1$ at the 5 percent level. If the 95 percent confidence interval of $(\beta_\pi - 1)$ includes zero, the computed inflation target $\pi^* = -\alpha/(\beta_\pi - 1)$ does not have a bounded confidence interval.

(ii) The 95 percent confidence interval of the inflation target should include some values between 1.5 percent and 2.0 percent. If this is not the case, we conclude that the estimated reaction function is not consistent with the definition of price stability and therefore it is not a good description of euro-area monetary policy. However, we do not (automatically) exclude models with a point estimate of π^* below 1.5 percent or above 2.0 percent, as long as the 95 percent confidence interval includes some values between 1.5 percent and 2.0 percent.

(iii) We require that the estimated parameter for projected inflation should be larger than that for projected real GDP growth: $\beta_\pi > \beta_y > 0$.[24]

We end up with 750 different specifications altogether, 13 of which meet the selection criteria described above. Typically, these specifications include the one-year-ahead inflation projection and

vintages. Then to instrument for the output growth gap (output growth nowcast, or one- to four-quarters-ahead forecast – potential output growth proxy), we use output growth nowcasts or forecasts from past data vintages (so that both elements entering the instrument for the output growth gap are from the same data vintage). Furthermore, we also use inflation nowcasts or forecasts from past data vintages as instruments. To be more specific, if output (inflation) forecast j = 0, 1, 2, 3, 4 periods ahead appears in the reaction function, we use the same forecast horizon from previous data vintages as an instrument. Finally, we instrument the nominal interest rate and the natural rate proxy with their lags. For all variables, we have used lags 2–4. Using different lags would give similar results.

[24]This is natural in the context of the ECB, as it does not have a dual mandate like the Federal Reserve. Furthermore, even if the parameters we estimate are not structural, also in the structural model of the euro area used at the ECB (NAWM), the estimated reaction function has this property.

the nowcast or one-quarter-ahead real GDP gap projection.[25] In these specifications, the ECB reacts rather strongly to projected inflation (the point estimate of the coefficient of inflation forecast ranges between 2.8 and 5.4, depending on specification). The reaction to the GDP growth gap is considerably more muted (typically the point estimate is roughly 0.4 or 0.5).

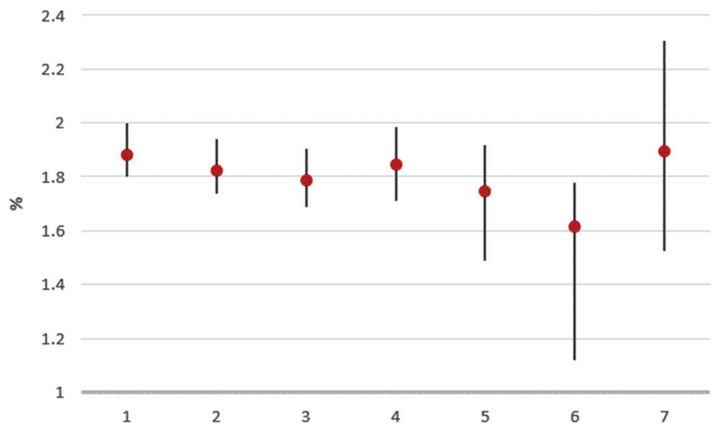
Figure 3 shows the computed inflation targets and their 95 percent confidence intervals based on sampling uncertainty from those seven model variants where the width of the confidence band is 100 basis points or narrower. In figure 3, the implied point estimate for the inflation target typically lies close to 1.8 percent. In the wider set of 13 specifications where the maximum width of the confidence band is 200 basis points, there are also a few rules with the inflation target at or below 1.6 percent. A rule with an inflation target of 2 percent or above is a rare exception. Furthermore, while the lower bound of the 95 percent confidence interval can be rather low in some rules, the upper bound typically lies below 2 percent. According to recursive estimations, the computed inflation target is relatively stable over time, apart from the period of the financial crisis.[26] In the specifications presented in figure 4, the point estimates across the models vary between 1.49 percent and 1.87 percent.

Finally, we augment equation (2) with the past inflation gap term. We allow the number of lags in the inflation gap term to vary from one to eight quarters. In this case, our preferred specification, in which all estimated coefficients are reasonable (i.e., of the expected sign and of meaningful size) includes the one-year-ahead inflation forecast, GDP growth nowcast, and a natural rate proxy based on the ex post real yield of 10-year German bunds. The monetary policy stance is measured by EONIA (average over the quarter) and the inflation gap is based on the past six quarters. The point estimate

[25] There are also five specifications which meet selection criteria (i) and (iii) but do not meet criterion (ii). All five specifications involve the two-quarters-ahead inflation projection. In these specifications the point estimate of the de facto inflation target is roughly 3 percent, while the lower bound of the 95 percent confidence band is typically at roughly 2.5 percent; the upper bound of the confidence band ranges from close to 4 percent to over 10 percent, depending on specification.

[26] Estimation results are available upon request.

Figure 3. Inflation Target: Point Estimates and 95 Percent Confidence Bands



Sources: ECB and authors' own calculations.
Notes: All specifications (1–7) displayed in the figure include the one-year-ahead inflation projection. Specifications 1–5 also include the GDP growth nowcast, while specification 6 includes the one-quarter-ahead, and specification 7 the two-quarters-ahead, GDP growth forecast. In specifications 1–6, the monetary policy stance is measured by EONIA (quarterly average), while in specification 7 the stance is measured by the MRO rate (end of period). The natural rate of interest is proxied by the real yield on German government bonds of different maturities: (1) five years (ex post), (2) three years (ex post), (3) two years (ex post), (4) five years (ex ante), (5) three years (ex ante), (6) one year (ex post), and (7) five years (ex ante).

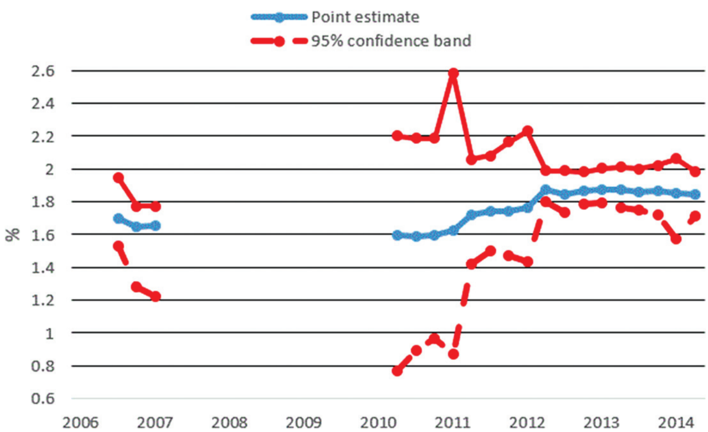
for the de facto inflation target is $\hat{\pi}^* = 1.77\%$ with a 95 percent confidence interval (capturing only sampling uncertainty) of 1.62–1.91 percent.[27]

To summarize, our estimations so far suggest that the ECB's monetary policy decisions are based on relatively short-term macroeconomic projections and the ECB's de facto inflation target lies between 1.7 percent and 1.8 percent. This finding is in line with the analysis of the inflation forecasts in the previous section.

It is useful to compare our results with survey-based measures of inflation expectations. Long-run inflation expectations in the ECB Survey of Professional Forecasters are more dispersed, but their

[27]Estimation results are available upon request.

Figure 4. Recursive Expanding Window Estimations



Sources: ECB and authors' own calculations.
Notes: The presented specification includes the four-quarters-ahead inflation projections, output growth nowcasts, and the ex ante real natural interest rate. The end of the estimation window is extended recursively from 2006:Q1 to 2014:Q2. Point estimates are not shown for periods for which the confidence band isn't bounded.

mean is comparable to our estimates of the de facto target. The distribution of long-term point estimates reveals that inflation expectations have been hovering between 1.7 and 2.0 percent during 2002–14. When looking at the aggregate probability distribution of long-term inflation expectations, the distribution is considerably wider than shown in figures 3 and 4. Even if most of the probability mass is between 1.5 and 1.9 percent, there is a considerable probability mass also between 0.5 and 1.4 percent and between 2.0 and 2.9 percent, and even beyond (see ECB 2019).

3.2 Reaction Functions where Cyclical Variable Is Output Growth

Measuring output gap and potential output in real time is notoriously difficult, and it is unlikely to be a good practice in policymaking to base policy on such uncertain measures of cyclical position of the economy. Indeed, the ECB does not discuss its output gap measures explicitly when it communicates its policy to the public.

Consequently, it is useful to consider reaction functions which do not directly rely on the output gap. The caveat is that the linear specification does not allow us to infer the de facto inflation target. However, with the inclusion of the nonlinear CL term, we can again indirectly infer the value of the de facto inflation target without a need to rely on an output gap measure.

3.2.1 Linear Forward-Looking Reaction Functions

For completeness, we discuss first the results from the linear specifications of the following form:

$$i_t = \rho i_{t-1} + (1-\rho)(\alpha + \beta_\pi(\pi^f_{t+j|t} - \pi^*) + \beta_y \triangle y^f_{t+k|t} + Dr^n_t). \quad (3)$$

In equation (3), i_t is the EONIA rate and the term $\pi^f_{t+j|t}$ refers to the ECB's projection of j-quarters-ahead HICP inflation and $\triangle y^f_{t+k|t}$ to the ECB's projection of k-quarters-ahead real GDP growth (instead of output growth gap as in equation (2)). Both ex ante and ex post proxies of the neutral real interest rate (r^n_t) based on the composite nominal yield of 10-year euro-area government bonds (see figure A.2 in appendix A) are considered; when the natural real rate enters (does not enter) into a reaction function, the dummy variable D is equal to one (zero). We set the inflation target to a number close to 2 percent, more specifically $\hat{\pi}^* = 1.9\%$.[28]

When estimating equation (3) with and without the natural real interest rate proxies, we again vary projection horizons from zero (nowcast) to four quarters.[29] We estimate 75 competing specifications altogether and choose the preferred specification following model selection criteria by which the estimated coefficients for forward-looking variables must imply that the interest rate reacts sufficiently strongly to projected inflation and output in order to stabilize the economy, and the estimated parameter for projected inflation should be larger than the one for projected real GDP growth.

[28] In the NAWM model of the ECB, the operational definition of price stability is also set at 1.9 percent (Christoffel, Coenen, and Warne 2008).

[29] We employ as instruments lagged variables from the same data vintage that is used in the monetary policy rule. We instrument inflation and output growth forecasts and nowcasts with lags 2–5 of (realized) inflation and output growth. As further instruments we use lags 3–4 of the nominal interest rates and lags 2–3 of the natural rate proxies. Using different lags would give similar results.

We also assess parameter stability as well as relevance of the real interest rate variable in the reaction function by running estimations in which we extend the pre-crisis sample (1999:Q4–2008:Q2) recursively quarter by quarter until the whole sample 1999:Q4–2014:Q2 is reached.

As in the previous section, the results support specifications with (i) very short-run (one-quarter-ahead) GDP growth projections; (ii) somewhat longer-term (one-year-ahead) inflation projections; and (iii) reaction functions including a proxy for the natural rate of interest.[30]

Table 1 summarizes our preferred linear specification, based on a four-quarters-ahead inflation gap and one-quarter-ahead output growth. According to this specification, the ECB reacts to a projected inflation gap about three times stronger than to a projected cyclical stance measured by output growth. The interest rate smoothing is rather high as expected and the relatively large coefficient for the inflation gap implies that the Taylor principle clearly holds: the real ex ante interest rate increases when inflation rises. Inclusion of a time-varying natural rate has only a small effect on the coefficient on output growth. The effect on the coefficient for expected inflation gap is somewhat larger, but this difference is partly mechanical, because we measure the real interest rate as a difference between a composite nominal yield of 10-year euro-area government bonds and real-time estimates of the current or one-period-ahead inflation forecast. Overall, it seems reasonable that the ECB conditions its interest rate decisions on the short end of the

[30]We obtain statistically significant coefficients also for the nowcast as well as one-quarter-ahead or four-quarters-ahead inflation, if the forecast horizon for real GDP growth is very short, i.e., zero (nowcast) or one quarter. Notably, a specification with the four-quarters-ahead inflation and one-quarter-ahead real GDP growth (i.e., $\pi^f_{t+4|t}$ and $\triangle y^f_{t+1|t}$) produces satisfactory coefficient estimates with either of the two proxies of the natural real interest rate, as well as without a natural rate proxy. Regarding parameter stability, we have estimated reaction functions with the four-quarters-ahead inflation ($\pi^f_{t+4|t}$) and one-quarter-ahead GDP ($\triangle y^f_{t+1|t}$) for the (pre-Lehman) period of 1999:Q4–2008:Q2, and then expanded the sample one quarter at a time until 2014:Q2. We obtain more stable coefficients for inflation and output growth with a natural interest rate proxy in the specification than without it. In addition, the specification using the ex ante natural real interest rate seems to work even better than the ex post natural real interest rate. All results are available upon request.

Table 1. Baseline Linear Reaction Function with Output Growth as Cyclical Variable

$i_t = \rho * i_{t-1} + (1-\rho) * (\alpha + \beta_\pi * (\pi^f_{t+4\|t} - 1.9) + \beta_y * \Delta y^f_{t+1\|t} + \tilde{r}^{10yr}_t)$				
	Coefficient	**Std. Error**	**t-statistic**	**Prob.**
ρ	0.84	0.044	19.23	0.0000
α	−0.95	0.737	−1.29	0.2049
β_π	4.45	0.832	5.34	0.0000
β_y	1.48	0.507	2.92	0.0057
J-statistic	6.07			
Prob(J-statistic)	0.73			

Notes: This table shows the GMM estimation results of our preferred linear reaction function of the ECB. The estimation sample is 1999:Q4–2014:Q2. See the main text for the definition of the variables and table B.1 (in appendix B) for alternative competing linear specifications. The reported J-statistic is the Sargent-Hansen test for validity of the instruments.

forecast horizon due to increasing difficulties to predict inflation and growth in the medium and longer term.

3.2.2 Symmetric Responses to Past Inflation Gaps

Next, we augment our preferred linear specification with a backward-looking "credibility loss term" CL_t so that

$$i_t = \rho i_{t-1} + (1-\rho)(\alpha + \beta_\pi(\pi^f_{t+4|t} - \pi^*) + \beta_y \triangle y^f_{t+1|t} + \gamma CL_t + r^n_t), \quad (4)$$

where the CL_t term is specified as in equation (1).

As discussed at the beginning of this section, this term captures the idea that the central bank may set the interest rate higher (lower) today if the past inflation gap is positive (negative) even if inflation is projected to be at the target in the near future.[31,32] Concerns for

[31]Monetary policy credibility measures proposed by de Mendonca and de Guimarães e Souza (2009) is also based on past deviations of inflation from the target.

[32]Using quite similar an approach, Dovern and Kenny (2017) investigate the impacts of "too low for too long" on long-term inflation expectations of professional forecasters in the euro area. They define an inflation "performance gap"

past inflation gaps may reflect, e.g., the central bank's desire and commitment to correct past errors. Note also that the credibility loss term weights large average past deviations of inflation from the target (π^*) more than small ones. Note that in Bernanke's (2017b) proposal of temporary price-level targeting, the key additional element in the policy rule is the term which captures the cumulative inflation shortfall since the beginning of zero lower bound until the exit date (see also Hebden and López-Salido 2018). The CL term captures a similar idea, but it introduces additional inertia in policymaking also at normal times when inflation deviates from the target.

When estimating equation (4), we allow for the length of the time span, i.e., the number of lags (q) over which the average past inflation is measured, to vary from one to eight quarters.[33] We also consider a number of different inflation targets π^*, at or below 2 percent; the lowest inflation target rate examined is chosen to be 1.6 percent in light of figure 1 and the results from section 3.1. This exercise allows us to draw additional indirect inference concerning both the ECB's de facto inflation target and the ECB's concerns of past inflation gaps.

Estimation results are summarized in table B.2 in appendix B. Based on our model evaluation criteria, longer credibility loss time spans, ranging up to six to eight lags, and lower inflation target rates (perhaps even as low as 1.6 percent or 1.7 percent) produce the most satisfactory and relatively robust coefficient estimates (estimated parameters seem to be relatively stable when the sample rolls recursively over the financial crisis toward 2014:Q2[34]). Our preferred nonlinear specification is reported in table 2. Compared with the linear specification in table 1, we now obtain smaller coefficients for interest rate smoothing and projected inflation gap while the ECB seems to react relatively strongly to past inflation gaps. The estimated output growth coefficient is roughly unchanged relative to the preferred linear specification.

as the difference between recent long-term inflation expectations and a moving average of past inflation rates.

[33]We do not instrument for the CL terms (which include past values of inflation), while the rest of the right-hand-side variables are instrumented as explained in section 3.2.1.

[34]The results are available upon request.

Table 2. Baseline Reaction Function with Symmetric Response to Past Inflation Gap

$$i_t = \rho * i_{t-1} + (1-\rho) * (\alpha + \beta_\pi * (\pi^f_{t+4|t} - 1.7) + \beta_y * \Delta y^f_{t+1|t} + \gamma * CL_t + \tilde{r}^{10yr}_t) \text{ where } CL_t = (\bar{\pi}_{t-1,t-7} - 1.7)|\bar{\pi}_{t-1,t-7} - 1.7|$$

	Coefficient	**Std. Error**	**t-statistic**	**Prob.**
ρ	0.77	0.051	15.30	0.0000
α	−1.51	0.396	−3.83	0.0004
β_π	3.61	0.798	4.53	0.0001
β_y	1.25	0.317	3.94	0.0013
γ	1.07	0.417	2.56	0.0145
J-statistic	6.58			
Prob(J-statistic)	0.68			

Notes: This table shows the GMM estimation results of our preferred reaction function of the ECB including symmetric reactions to past inflation gaps. The estimation sample is 1999:Q4–2014:Q2. See the main text for the definition of the variables and table B.2 (in appendix B) for alternative competing linear specifications. The reported J-statistic is the Sargent-Hansen test for validity of the instruments.

In sum, we find that a concern for past errors seems to have played a role in the ECB's policy decisions. Quite intuitively, however, the ECB has responded only to persistent inflation gaps as indicated by the long lags of the credibility loss term.[35] Consistently with the findings from section 3.1, these results also suggest that the ECB's de facto inflation target has been considerably below 2 percent, perhaps even as low as 1.6 percent or 1.7 percent.[36] Hence, the results for the de facto inflation target do not seem to be overly sensitive to the choice of the cyclical variable or even the inclusion of the past inflation gap term.[37]

[35] This is reasonable, since monetary policy is not expected to respond to temporary shocks to inflation such as large variations in energy prices.

[36] Bletzinger and Wieland (2017), using the ECB Survey of Professional Forecasters and European Commission estimates for potential growth, and considering a target range of 1.5–2.0 percent, conclude that the ECB point inflation target is 1.7 percent. Furthermore, when estimating first-difference policy rules by Orphanides (2003), Hartmann and Smets (2018) conclude that the ECB's implicit target is 1.81 percent.

[37] As a further robustness check, we have also estimated a number of linear reaction functions without any real activity measure. These specifications are based

3.2.3 Asymmetric Responses to Past Inflation Gaps

Next, we consider possible asymmetry in the ECB's policymaking, i.e., we allow for different responses to positive and negative past inflation gaps. We estimate the following specification:

$$i_t = \rho * i_{t-1} + (1-\rho) * (\alpha + \beta_\pi * (\pi^f_{t+4|t} - \pi^*) + \beta_y * \triangle y^f_{t+1|t} + \gamma_1 * CL^+_t + \gamma_2 * CL^-_t + r^n_t), \quad (5)$$

where

$$CL^+_t = D^{CL}_t * CL_t$$
$$CL^-_t = (1 - D^{CL}_t) * CL_t.$$

In equation (5), the dummy variable D^{CL}_t is equal to one (zero) if $CL_t > 0$ ($CL_t < 0$). The coefficient γ_1 captures monetary policy reactions to past positive inflation gaps, and the coefficient γ_2 to past negative inflation gaps. In order to measure the ECB's credibility concerns in a meaningful way, the parameters γ_1 and γ_2 must be positive, but their sizes may differ.

Again, we run several competing specifications in order to draw some inference concerning both the ECB's de facto inflation target and the ECB's concerns of past inflation gaps. In table B.3 in appendix B, the credibility loss term is based on one to eight lags of actual inflation and the inflation target varies from 1.6 to 2.0 percent. Consistent with our results for symmetric reaction functions, table B.3 in appendix B indicates that the time span of the past inflation gap should be rather long, ranging from six to eight quarters (the ECB reacts only to rather persistent inflation gaps). However, as our preferred specification in table 3 reveals, now the inflation target closer to 2 percent seems more appropriate but the ECB's policy is asymmetric: it responds more aggressively to positive than to negative inflation gaps (i.e., the parameter estimate for γ_1 is significantly larger than for γ_2).[38] Such an asymmetric

on the assumption that the ECB policy responds only to projected inflation. The results are not sensitive to the inclusion or exclusion of a cyclical variable.

[38]Note that a time-invariant potential output can be calculated as $\triangle y^* = (\pi^* - \alpha)/\beta_y$. According to the symmetric reaction function (asymmetric reaction function), the average projected potential output growth is 2.5 percent

Table 3. Baseline Reaction Function with Asymmetric Response to Past Inflation Gap

$i_t = \rho * i_{t-1} + (1-\rho) * (\alpha + \beta_\pi * (\pi^f_{t+4\|t} - 2.0) + \beta_y * \Delta y^f_{t+1\|t} + \gamma_1 * CL^+_t + \gamma_2 * CL^-_t + r^n_t)$ where $CL_t = (\bar{\pi}_{t-1,t-7} - 2.0)\|\bar{\pi}_{t-1,t-7} - 2.0\|$				
	Coefficient	**Std. Error**	**t-statistic**	**Prob.**
ρ	0.79	0.046	16.94	0.0000
α	−1.92	0.755	−2.54	0.0153
β_π	1.23	0.585	2.10	0.0424
β_y	1.69	0.379	4.47	0.0001
γ_1	8.00	2.606	3.07	0.0039
γ_2	0.63	0.312	2.01	0.0518
J-statistic	6.66	F-statistic ($\gamma_1 = \gamma_2$)		6.84
Prob(J-statistic)	0.67	Prob(F-statistic)		0.013

Notes: This table shows the GMM estimation results of our preferred asymmetric reaction function of the ECB. The estimation sample is 1999:Q4–2014:Q2. See the main text for the definition of the variables and table B.3 (in appendix B) for alternative competing linear specifications. The reported J-statistic is the Sargent-Hansen test for validity of the instruments. The F-statistic is obtained from the test for asymmetry of the reaction function, by testing equality of the positive and negative credibility loss term coefficient estimates.

reaction to past inflation gaps implies that, over a long period of time, inflation will be below 2 percent, i.e., asymmetry itself lowers the de facto inflation target.[39]

As for the estimated coefficient for inflation, this policy rule suggests that the ECB is considerably backward looking. The estimated coefficient for the one-year-ahead inflation forecast is small, three to four times smaller than in the previous estimations (see tables 1

(2.3 percent). These numbers naturally deviate somewhat from the ex post data, reflecting both real-time uncertainty of future real GDP growth and end-point problems. At the same time, the implied projected potential growth rates are in line with the Eurosystem/ECB staff real GDP growth projections. The projected growth rates at the end of each projection horizon are good proxies for the real-time estimates of the projected potential output growth. As already discussed in section 2, the ECB's projections of real GDP growth converge to values between slightly below 2 percent and 3 percent.

[39]According to the asymmetry analysis by Hartmann and Smets (2018), the ECB's accommodative policy responses are mainly due to decreasing output growth projections, and tightening policy responses are mainly due to above the target inflation projections.

and 2). At the same time, the point estimate for the output growth term is somewhat higher.

In order to assess whether asymmetric policy responses reflect the zero lower bound of interest rates, we have reestimated the reaction function (5) using a shadow rate instead of EONIA for a longer sample 1999:Q1–2016;Q4. The results are qualitatively unchanged.[40]

In summary, the ECB's definition of price stability seems to manifest itself in two alternative ways. Either the de facto target of the ECB is significantly below 2 percent and policy responses to past inflation gaps are symmetric, or the ECB's inflation target is close to 2 percent and it reacts more strongly to past positive than to past negative inflation gaps. The two policy rule specifications also have other interesting differences: in the case of the symmetric specification, the policy response to the projected inflation is clearly higher (three times higher) and the response to past inflation gaps substantially lower than in the case of the asymmetric specification. While in both cases a reaction to past inflation gaps implies that the ECB attempts to correct past inflation misses, this behavior is particularly strong under the asymmetric specification. Given that there is no substantial difference in the interest rate smoothing coefficient, the ECB appears to be more forward looking under the symmetric specification. As for now, there is no clear statistical criteria by which we could give preference to either of the two reaction function specifications.

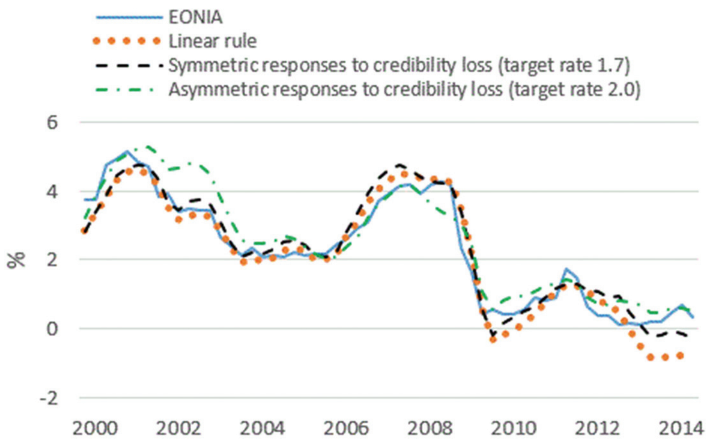
4. Predictive Performance of Different Reaction Functions

4.1 In-Sample Predictions

The performance of our preferred reaction functions from tables 1 to 3 can be assessed by comparing their in-sample predictions with the EONIA interest rate. Figure 5 indicates that the asymmetric reaction function deviates at times significantly from the EONIA rate and from predictions of the two other functions, especially at the beginning of the sample, when the euro-area inflation was quite often above 2 percent. During 2005–07, however, the asymmetric

[40] Estimation results are available upon request.

Figure 5. Dynamic In-Sample Predictions of Different Reaction Functions



Sources: ECB, Thomson Reuters, and authors' own calculations.
Note: The dynamic in-sample predictions are based on our preferred specifications of the ECB's reaction function reported in sections 3.2.1–3.2.3.

reaction function tracks relatively well the EONIA rate. In mid-2008 it misses the increase in the EONIA rate, and from there on it stays most of the time above EONIA and also above predictions of the two other reaction functions. Both the symmetric and linear reaction function would have implied a stronger interest rate hike prior to the financial crisis, but in general more lax policy after 2009.

The linear reaction function, which only responds to the expected future path of the economy and not at all to past inflation gaps, generates the lowest interest rate path (i.e., the most accommodative monetary policy stance) at the end of the sample. This reflects relatively strong responses to the projected slowdown of inflation during this period. The symmetric nonlinear reaction function with a low de facto target inflation generates a similar path but yields a somewhat less accommodative policy stance, because it puts weight on a past positive inflation gap (see figure A.3 in appendix A) and the impact of the projected slowdown of inflation is smaller. Excluding the end of the sample, the linear and symmetric nonlinear reaction functions give rather similar predictions for the interest rate path until about 2012. According to these specifications, the zero lower bound would have been reached in 2009, i.e., much earlier than it was actually reached. Instead, according to the asymmetric nonlinear reaction function the zero lower bound would not have been reached at all.

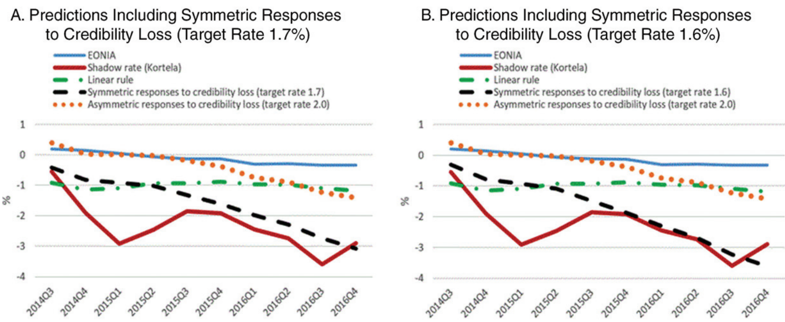
4.2 Out-of-Sample Predictions and Comparison with Shadow Rates

How do the estimated reaction functions describe the monetary policy stance under unconventional monetary policy measures when the interest rate has hit the zero lower bound? In other words, are unconventional and conventional measures determined by the same basic principles, so that unconventional measures can be thought of as a continuation of conventional monetary policy when the zero lower bound is reached? Assuming that one of our preferred policy rule specifications provides a reasonable description of the ECB monetary policy until 2014, the same policy strategy should have been applied also afterwards for this policy to be time consistent. In fact, the ECB has recently emphasized in its communication that the reaction function has not changed despite the zero lower bound period and the instruments of monetary policy (Hutchinson and Smets 2017).

In July 2013, the ECB introduced explicit forward guidance to inform markets and the public on its future intentions with regard to key policy rates, and in January 2015 the ECB launched an expanded asset purchase program (APP) to address risks of too-prolonged a period of low inflation. Previously, the ECB had emphasized in its communication that it does not pre-commit on monetary policy decisions. Also, other unconventional monetary policy measures were adopted in 2015 and 2016 in order to maintain an accommodative stance of monetary policy. The deposit rate was cut in June and September 2014 to −0.2 percent, reaching for the first time a negative territory.[41]

[41] Since mid-2014, in order to provide financing to euro-area credit institutions, two series of targeted longer-term refinancing operations (TLTROs) were introduced: the first series of eight operations (LTRO-I) was announced in June 2014, and a second series of four operations (LTRO-II) in March 2016. In September 2014, the ECB made announcements of the third covered bond purchase program (CBPP3), an asset-backed securities purchase program (ABSPP), and a further deposit facility rate cut. APP purchases were started in March 2015 and they

Figure 6. Shadow Rate by Kortela (2016) and Predictions Based on Different Reaction Functions



Sources: ECB, authors' own calculations, and Kortela (2016) for the shadow rate.
Notes: The symmetric responses to a credibility loss refer to a reaction function with a low de facto inflation target (1.7 percent or 1.6 percent). The asymmetric responses to a credibility loss refer to a reaction function with an inflation target of 2.0 percent.

To analyze the recent euro-area monetary policy stance, we use our preferred reaction functions (estimated for the period 1999:Q4–2014:Q2) to produce dynamic out-of-sample forecasts for the period 2014:Q3–2016:Q4. Our aim is to assess how closely the whole path of dynamically predicted interest rate matches to the measure of the ECB's monetary policy stance. In each quarter, the prediction is conditional on the Eurosystem/ECB staff real-time forecasts and the lagged prediction of the interest rate from the corresponding rule.

In figure 6A, the implied interest rates are compared with the shadow rate estimated by Kortela (2016). He argues that the euro-area shadow rate had gradually decreased to about −3 percent by the end of 2016, while a temporary increase was experienced in 2015. The dynamic out-of-sample forecast of the interest rate implied

---

were recalibrated in December 2015, March 2016, and December 2016. The ECB took a number of nonstandard measures already in the earlier phase of the crisis, but these measures were mainly targeted to provide ample liquidity for the euro-area banks and they were taken in tandem with standard interest rate cuts.

by our linear rule remains negative and stable around −1 percent throughout the whole zero lower bound period; it is roughly 1 percentage point below the EONIA rate but considerably higher than the shadow rate for most of the period.

The nonlinear reaction functions taking into account a credibility loss imply falling interest rates over the period 2014:Q3–2016:Q4. The symmetric nonlinear reaction function with a low de facto inflation target of 1.7 percent seems to track the shadow rate considerably better than the asymmetric nonlinear reaction function with an inflation target of 2 percent. This suggests, tentatively, that the ECB's definition of price stability is best characterized by an inflation target that is markedly below 2 percent, but the ECB is symmetric in its reactions to past inflation gaps. If we consider the symmetric reaction function based on a lower inflation target of 1.6 percent, which is also a plausible target rate according to our estimation results shown in table B.2 in appendix B, the implied predictions are even closer to the shadow rate (figure 6B).

Finally, as a robustness check, we compare the same out-of-sample predictions to another shadow interest rate estimated by Wu and Xia (2016).[42] As shown in appendix C, their shadow rate is steadily decreasing to about −5 percent in 2016. Compared with Kortela's (2016), their analysis indicates even more accommodative a monetary policy stance in the euro area in recent years. The main technical difference between Kortela and Wu and Xia is that Kortela allows for a time-varying lower bound for the euro area, reflecting the expected path of the deposit facility rate. The Wu and Xia methodology is based on a constant lower bound assumption. Nevertheless, also in Wu and Xia's case, the symmetric reaction functions with a low de facto inflation target (1.6 percent or 1.7 percent) seem to characterize most accurately the conducted policy in the euro area. In general, estimated shadow rates are of course subject to large uncertainty. As reported in Hartmann and Smets (2018), the shadow rates vary between close to −8 percent and 0 percent in the period 2014–17.

[42]The shadow interest rate constructed by Wu and Xia (2016) is based on an analytical representation for bond prices in an SRTSM model (https://www.quandl.com/data/SHADOWS/EUROPE-European-Central-Bank-Shadow-Rate).

5. Conclusions

In the recent discussion of the ECB's monetary policy, there has been a vivid debate on what the ECB's de facto target is, whether the reaction function is symmetric, and, if it is, around which inflation target. Increased clarity about the central bank's reaction function and inflation target helps to anchor inflation expectations by reducing uncertainty about the central bank's future actions.[43] In this paper, we have shed some more light on the ECB's de facto inflation target, possible asymmetry of policy responses, and its reaction function in general. To do this, we have estimated a large number of competing specifications for the ECB's reaction function. We have used extensively the real-time projections from the Eurosystem/ECB staff macroeconomic projection exercises conducted in 1999:Q4–2016:Q4. A quarterly real-time data set has enabled us to assess realistically the ECB Governing Council's monetary policy decisionmaking by estimating the reaction functions with the same information it has available when it decides on the monetary policy stance. After estimating reaction functions including different levels of the de facto inflation target, different cyclical variables, a time-varying natural rate, varying degrees of backward-looking and forward-looking information contained in our real-time data, and asymmetry, we have arrived at the following robust findings.

First, the de facto inflation target of the ECB is well below 2 percent, perhaps even as low as 1.6–1.8 percent. This finding is also consistent with the fact that the Eurosystem/ECB staff medium-term inflation projections have had a tendency to converge rapidly on values well below 2 percent.

Second, the reaction function specifications which include both forward-looking information from the projections and past inflation developments seem to characterize best the ECB's monetary policy decisions during the whole sample. We find some evidence on asymmetry around the presumed 2 percent inflation target, but the dynamic out-of-sample predictions of the symmetric reaction

[43]See, e.g., Taylor (2018) for general discussion on the benefits of rule-based policy and Bernanke and Mishkin (1997) and Bundick and Smith (2018) for the importance of a specific numerical inflation target for anchoring inflation expectations in the United States.

function with a 1.6–1.7 percent de facto target is better in line with the evolution of shadow rates than the asymmetric reaction function during the zero lower bound period.

Our results suggest that, in general, the ECB’s policy reaction function follows the basic optimality principles in accordance with its mandate. We find that the ECB Governing Council responds relatively strongly to the expected short-term future course of the economy—i.e., its forecasts for inflation and the measure of real economic activity—but it also aims at correcting past persistent deviations of inflation from the target. The forward-looking nature of its policy is motivated by the fact that monetary policy affects the economy only gradually, hence inflation and output forecasts should be an integral part of the inflation-targeting strategy. That policy has also a backward-looking element, consistent with the ECB communication, according to which the launch of asset purchase programs and other unconventional policy measures can be justified as a response to too-prolonged a period of low inflation.[44] At the same time, however, asymmetric responses to inflation and/or a low de facto inflation target indicated by our findings may hamper the ECB’s ability to achieve its inflation aim. There are a number of reasons for this.

Firstly, when approaching the inflation target from below, the central bank may need to tolerate inflation rates above the target. Overshooting the target for a limited time may help the central bank to achieve its inflation aim faster and more efficiently when interest rates are at the zero lower bound. Under credible monetary policy, overshooting the target raises inflation expectations and lowers the ex ante real interest rate. This boosts consumption and investment and therefore reduces economic slack in the standard New Keynesian type of models.

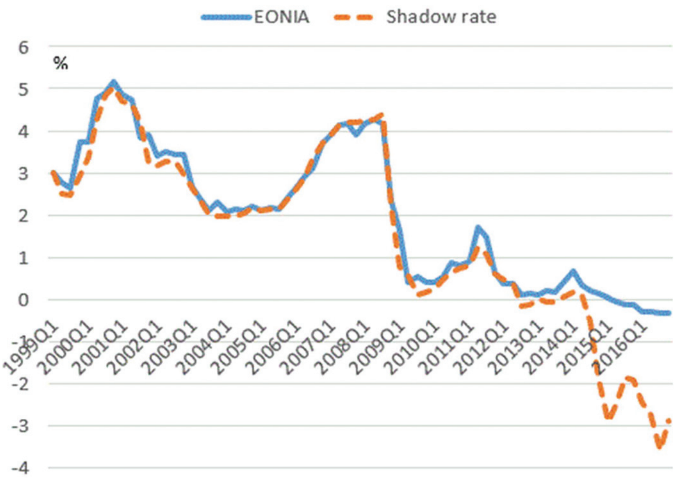
Secondly, for a given equilibrium real interest rate, anchoring of inflation expectations to a relatively low level also leads to low nominal rates over the business cycle. This reduces the scope to absorb shocks in economic downturns and increases the likelihood of hitting the zero lower bound. With forward-looking price-setting behavior,

[44] See http://www.ecb.europa.eu/press/key/date/2018/html/ecb.sp180314_1.en.html.

the expectation that monetary policy has less scope to absorb negative shocks in the future can further lower the current inflation. Miles et al. (2017) have also recently stressed that in the current low inflation environment, overshooting the target is necessary and the targeted rate of inflation should not be too low.[45]

Appendix A. Additional Figures and Tables

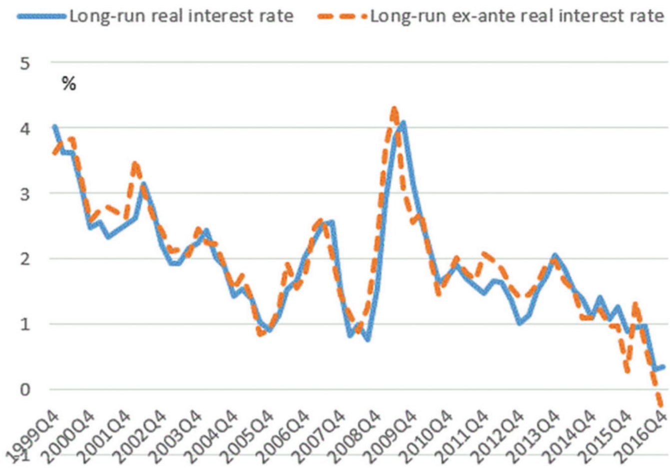
Figure A.1. EONIA and Shadow Rate at End of Each Quarter



Sources: Thomson Reuters (EONIA) and Kortela (2016) (the shadow rate).

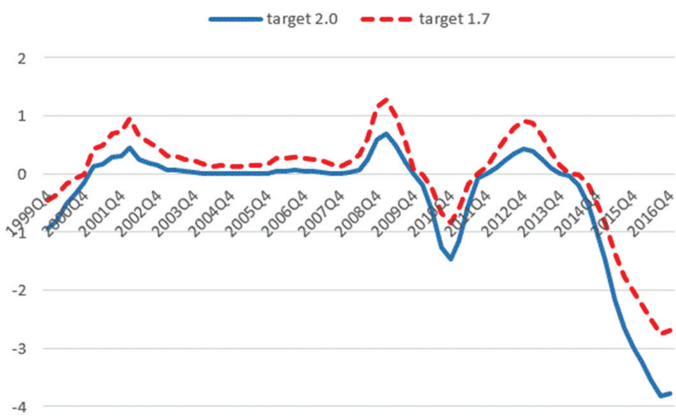
[45]In the United States, too, the level and symmetry of the targeted rate of inflation has been discussed recently. See, for example, speeches by Evans: https://www.chicagofed.org/publications/speeches/2017/11-15-2017-low-inflation-and-symmetry-of-two-percent-target-charles-evans-london-ubs, https://www.chicagofed.org/publications/speeches/2017/09-25-17-puzzle-low-inflation-implications-monetary-policy.

Figure A.2. Proxies of Long-Run Natural Real Interest Rate



Sources: ECB, Thomson Reuters, and authors’ own calculations.
Note: The long-run real interest rate is equal to the difference of a euro-area composite nominal yield of 10-year government bonds and the real-time nowcast or one-quarter-ahead forecast of inflation rate.

Figure A.3. Values of Credibility Loss Term



Sources: ECB and authors’ own calculations.
Note: The horizon over which the average inflation is measured is seven quarters ($\bar{\pi}_{t-1,t-7}$). See the main text for the definition of the credibility term.

Table A.1. Characteristics of Projection Data

	$\Delta\pi^f_{t\|t}$	$\Delta\pi^f_{t+4\|t}$	$\Delta\pi^f_{t+8\|t}$	$\Delta\pi^f_{t+11\|t}$	**HICP Change**
Mean	1.78	1.60	1.64	1.76	1.75
Standard Dev.	0.97	0.34	0.22	0.20	0.98
Sample Size	69	69	68	17	69
	$\Delta y^f_{t\|t}$	$\Delta y^f_{t+4\|t}$	$\Delta y^f_{t+8\|t}$	$\Delta y^f_{t+11\|t}$	**GDP Growth**
Mean	1.10	1.67	2.05	2.12	1.33
Standard Dev.	1.57	0.81	0.48	0.49	1.99
Sample Size	69	69	68	17	69

Sources: ECB and authors' own calculations.
Note: The sample spans from 1999:Q4 to 2016:Q4 (69 quarters in total).

Appendix B. Summary of Estimations

Table B.1. Coefficients of Inflation and GDP Growth in Reaction Function (1), with Different Projection Horizons for Inflation (rows) and Output Growth (columns)

	$\Delta y^f_{(t\|t)}$		$\Delta y^f_{(t+1\|t)}$		$\Delta y^f_{(t+2\|t)}$		$\Delta y^f_{(t+3\|t)}$		$\Delta y^f_{(t+4\|t)}$	
A. Linear Policy Reaction Function without a Natural Rate of Interest										
$\pi^f_{(t\|t)}$	−0.50	1.30*	0.88*	1.46*	0.02	3.76*	0.00	2.89*	0.06	3.08*
$\pi^f_{(t+1\|t)}$	0.87*	1.06*	1.05*	1.21*	0.37	3.36*	0.11	2.84*	0.63	2.98*
$\pi^f_{(t+2\|t)}$	0.86	1.14*	0.76	3.19*	0.28	3.55*	0.08	2.86*	0.33	3.07*
$\pi^f_{(t+3\|t)}$	0.36	1.28*	0.00	3.94*	−0.47	4.07*	−0.11	2.94*	0.05	3.07*
$\pi^f_{(t+4\|t)}$	**2.34***	**0.94***	**2.97***	**1.84***	1.85	2.69*	0.75	2.59*	1.26	2.75*
B. Linear Reaction Function with r^{10yr}_t as a Proxy for the Natural Rate of Interest										
$\pi^f_{(t\|t)}$	−12.99	8.01	1.78*	5.39	0.04	4.95*	0.65	3.05*	0.50	2.79*
$\pi^f_{(t+1\|t)}$	**2.22***	**0.76***	2.31	6.23	0.20	4.82*	0.77	3.34*	0.66	2.63*
$\pi^f_{(t+2\|t)}$	**3.72***	**0.85***	2.76	5.60	0.78	4.37*	1.22	3.15*	1.00	2.63*
$\pi^f_{(t+3\|t)}$	2.26	0.62	3.69	5.47	0.25	4.82*	1.31	2.89*	1.40	2.55*
$\pi^f_{(t+4\|t)}$	**4.51***	**0.73***	**4.82***	**1.19***	**4.51***	**1.48***	**4.46***	**1.20***	**4.48***	**1.09***
C. Linear Reaction Function with $\tilde{r}^{10yr}_t$ as a Proxy for the Natural Rate of Interest										
$\pi^f_{(t\|t)}$	−4.32	3.35	1.16	7.49	0.13	4.41*	0.28	2.98*	1.72*	3.56*
$\pi^f_{(t+1\|t)}$	0.21	1.26	1.61	7.72	0.36	4.29*	0.52*	2.72*	1.05*	2.97*
$\pi^f_{(t+2\|t)}$	0.49	1.51	1.41	7.88	0.35	4.26	0.62	2.71	3.14*	5.12
$\pi^f_{(t+3\|t)}$	−4.34	2.91	−5.92	27.86	−0.59	4.96	0.40	2.91	3.68	6.67
$\pi^f_{(t+4\|t)}$	3.84*	0.84	**4.45***	**1.48***	**3.19***	**2.32***	**3.21***	**1.72***	**3.12***	**1.95**

Notes: 1. For each pair of numbers, the first entry is the coefficient of inflation, β_π, while the second entry is the coefficient of real GDP growth, β_y. 2. Coefficient estimates which are statistically significant, at least at the 5 percent level, are marked by *. 3. Bolded numbers mark model variants, where (i) both coefficients β_π and β_y are statistically significant, (ii) the coefficient of inflation is greater than one, and (iii) the coefficient of inflation is greater than the coefficient of real GDP growth. 4. We have added a gray background color to the combinations of inflation and output projection horizons ($\pi^f_{t+4|t}, \Delta y^f_{t+1|t}$) which satisfy the criteria (i)–(iii) in all the reaction functions (panels A, B, and C), with and without a natural rate proxy.

Table B.2. Symmetric Monetary Policy Responses to Credibility Loss in Reaction Function (2)

	Target 1.6	Target 1.7	Target 1.8	Target 1.9	Target 2.0
A. Reaction Functions that Include the Long Real Interest Rate r_t^{10yr} as a Proxy for the Natural Rate					
$\bar{\pi}_{t-1,t-1}$	−0.61	−0.65	−0.68	−0.70	−0.61
$\bar{\pi}_{t-1,t-2}$	−0.49	−0.40	−0.20	−0.14	−0.13
$\bar{\pi}_{t-1,t-3}$	−0.11	−0.10	−0.11	−0.13	−0.17
$\bar{\pi}_{t-1,t-4}$	0.11	0.11	0.10	0.08	0.02
$\bar{\pi}_{t-1,t-5}$	0.94*	0.70*	0.58*	0.50*	0.44*
$\bar{\pi}_{t-1,t-6}$	1.21*	1.02*	0.82*	0.64*	0.09
$\bar{\pi}_{t-1,t-7}$	**1.50***	**0.94***	0.24	0.19	0.53
$\bar{\pi}_{t-1,t-8}$	**2.90***	0.77	0.42	0.39	0.53
B. Reaction Functions that Include the Ex Ante Long Real Interest Rate $\tilde{r}_t^{10yr}$ as a Proxy for the Natural Rate					
$\bar{\pi}_{t-1,t-1}$	−0.53	−0.44	−0.07	−0.02	0.01
$\bar{\pi}_{t-1,t-2}$	0.20	0.16	0.13	0.11	0.09
$\bar{\pi}_{t-1,t-3}$	0.06	0.06	0.06	0.06	0.05
$\bar{\pi}_{t-1,t-4}$	−0.43	−0.10	0.02	0.05	0.06
$\bar{\pi}_{t-1,t-5}$	−0.44	−0.44	−0.24	−0.01	0.08
$\bar{\pi}_{t-1,t-6}$	0.29	0.29	0.31	0.31	0.29
$\bar{\pi}_{t-1,t-7}$	**1.42***	**1.07***	0.76	0.62	0.58
$\bar{\pi}_{t-1,t-8}$	3.20*	1.30	1.19	1.16	0.96

Notes: 1. The table reports estimates of the coefficient γ for the credibility loss term CL_t, for different spans of past inflation (rows) and inflation targets (columns). 2. Coefficient estimates γ which are of the correct sign (positive) and statistically significant, at least at the 5 percent level, are marked by *. 3. We have bolded the model specifications where also the coefficients of inflation and GDP growth projections (β_π and β_y, not shown in the table) are positive and statistically significant, and in addition $\beta_\pi > \beta_y$. 4. We have added a gray background color to the combinations of past inflation spans and inflation targets which meet the conditions 2 and 3 in both types of reaction functions considered here (i.e., these combinations are bolded in both panels of the table, A and B).

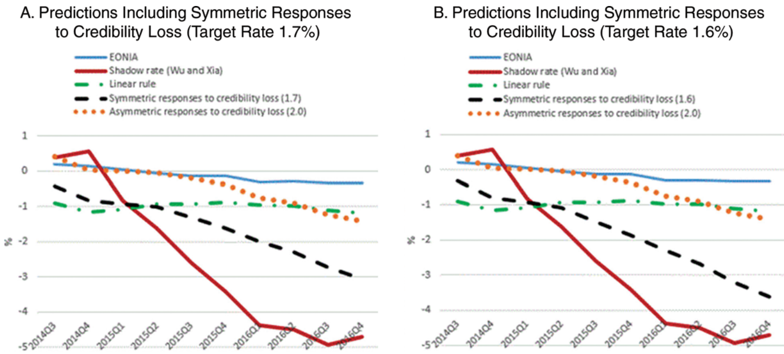
Table B.3. Asymmetric Monetary Policy Responses to Credibility Loss in Reaction Function (3)

	Target 1.6		**Target 1.7**		**Target 1.8**		**Target 1.9**		**Target 2.0**	
A. Reaction Functions that Include the Long Real Interest Rate r_t^{10yr} as a Proxy for the Natural Rate										
$\bar{\pi}_{t-1,t-1}$	0.60	0.36*	0.61	0.26	0.62	0.15	0.63	0.03	−1.55	−0.35
$\bar{\pi}_{t-1,t-2}$	0.68	0.45	0.68	0.31	0.60	0.14	−1.24	−0.08	−1.42	−0.10
$\bar{\pi}_{t-1,t-3}$	−0.91	0.25	−1.00	0.15	−1.03	0.07	−0.47	−0.04	−0.29	−0.15
$\bar{\pi}_{t-1,t-4}$	−2.31	0.61*	−2.17	0.44*	−2.19	0.34	0.49	−0.01	0.02	0.00
$\bar{\pi}_{t-1,t-5}$	−3.17	1.09*	−3.83	0.82*	−4.37	0.62*	−4.80	0.48*	−5.32	0.39*
$\bar{\pi}_{t-1,t-6}$	2.61*	0.95	2.96*	0.97*	−0.51	0.83*	−1.06	0.68*	−1.97	0.57*
$\bar{\pi}_{t-1,t-7}$	2.89*	1.00	3.67*	0.77	**4.73***	**0.64**	**6.14***	**0.61**	**8.00***	**0.63***
$\bar{\pi}_{t-1,t-8}$	3.30*	−1.02	4.18*	−1.08	5.66*	−1.21	8.39*	−1.19	12.77*	−0.71
B. Reaction Functions that Include the Ex Ante Long Real Interest Rate $\tilde{r}_t^{10yr}$ as a Proxy for the Natural Rate										
$\bar{\pi}_{t-1,t-1}$	0.43	0.51*	0.42	0.44*	0.41	0.37*	0.38	0.31*	0.34	0.24*
$\bar{\pi}_{t-1,t-2}$	0.65	0.54*	0.68	0.48*	0.71	0.41*	0.71	0.34	0.68	0.27
$\bar{\pi}_{t-1,t-3}$	−1.66	0.72*	−1.74	0.57*	−1.81	0.45*	2.75	0.24	−0.59	0.09
$\bar{\pi}_{t-1,t-4}$	−1.24	0.53*	−1.31	0.43*	−1.43	0.36*	−1.62	0.30*	−1.91	0.25*
$\bar{\pi}_{t-1,t-5}$	−1.86	0.96*	−2.08	0.76*	−2.35	0.61*	−2.74	0.50*	−3.34	0.42*
$\bar{\pi}_{t-1,t-6}$	0.74	0.26	0.37	0.40	−2.62	1.15*	−2.98	0.91*	−3.55	0.74*
$\bar{\pi}_{t-1,t-7}$	2.4*	0.64	3.06*	0.43	3.88*	0.31	4.92*	0.28	**6.21***	**0.31**
$\bar{\pi}_{t-1,t-8}$	3.65*	−8.67	4.57*	−4.26	5.83*	−1.85	7.46*	−0.77	9.69*	−0.14

Notes: 1. The table reports estimates of monetary policy reactions to a positive past inflation gap (coefficient γ_1, left entry in each pair of numbers) and a negative past inflation gap (coefficient γ_2, right entry in each pair of numbers), for different spans of past inflation (rows) and different de facto inflation targets (columns). 2. Coefficient estimates which are of the correct sign (positive) and statistically significant, at least at the 5 percent level, are marked by *. 3. Bolded numbers mark model variants where (i) both coefficients are of the correct sign (positive), (ii) at least the reaction to a past positive inflation gap (γ_1) is significantly different from zero, and (iii) the policy reaction to past positive deviations from the inflation target is significantly stronger than the reaction to past negative deviations from the target (i.e., γ_1 is significantly larger than γ_2, at least at the 5 percent level). 4. We have added a gray background color to the combinations of past inflation span and inflation target ($\bar{\pi}_{t-1,t-7}$ and target 2.0), which satisfy the criteria (i), (ii), and (iii) in both types of reaction functions considered here (i.e., in panels A and B of the table).

Appendix C. Alternative Shadow Rate by Wu and Xia (2016)

Figure C.1. Predictions Based on Different Reaction Functions



Sources: ECB, authors' own calculations, and Wu and Xia (2016) for the shadow rate.
Note: The symmetric responses to a credibility loss refer to a reaction function with a low de facto inflation target (1.6 percent or 1.7 percent). The asymmetric responses to a credibility loss refer to a reaction function with an inflation target of 2.0 percent.

References

Aguiar, A., and M. M. F. Martins. 2008. "Testing for Asymmetries in the Preferences of the Euro-Area Monetary Policymaker." *Applied Economics* 40 (13): 1651–67.

Alessi, L., E. Ghysels, L. Onorante, R. Peach, and S. Potter. 2014. "Central Bank Macroeconomic Forecasting During the Global Financial Crisis: The European Central Bank and Federal Reserve Bank of New York Experiences." *Journal of Business and Economic Statistics* 32 (4): 483–500.

Apel, M., and C. A. Claussen. 2017. "Inflation Targets and Intervals — An Overview of the Issues." *Economic Review* (Sveriges Riksbank) 2017 (1): 83–103.

Belke, A., and J. Klose. 2011. "Does the ECB Rely on a Taylor Rule During the Financial Crisis? Comparing Ex-post and Real Time

Data with Real Time Forecasts." *Economic Analysis and Policy* 41 (2): 147–71.

Bernanke, B. 2017a. "Monetary Policy in a Low Interest World: Comment." In *Brookings Papers on Economic Activity* (Spring): 380–89.

———. 2017b. "Monetary Policy in a New Era." Speech at the Peterson Institute Conference on Rethinking Macroeconomic Policy, Washington, DC, October 12–13.

Bernanke, B., and F. Mishkin. 1997. "Inflation Targeting: A New Framework for Monetary Policy?" *Journal of Economic Perspectives* 11 (2): 97–116.

Bletzinger, T., and V. Wieland. 2017. "Lower for Longer: The Case of the ECB." *Economics Letters* 159 (October): 123–27.

Bullard, J. 2018. "A Primer on Price Level Targeting in the U.S." Speech at the Federal Reserve Bank of St. Louis, St. Louis, Missouri, January 10.

Bundick, B., and A. L. Smith. 2018. "Did Communicating a Numerical Inflation Target Anchor U.S. Inflation Expectations?" *The Macro Bulletin* (Federal Reserve Bank of Kansas City) January 17: 1–4.

Charemza, W., and D. Ladley. 2016. "Central Banks' Forecasts and Their Bias: Evidence, Effects and Explanation." *International Journal of Forecasting* 32 (3): 804–17.

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." ECB Working Paper No. 944 (October).

Clarida, R. 2012. "What Has — and Has Not — Been Learned about Monetary Policy in a Low-Inflation Environment? A Review of the 2000s." *Journal of Money, Credit and Banking* 44 (s1): 123–40.

Clarida, R., J. Galí, and M. Gertler. 1998. "Monetary Policy Rules in Practice: Some International Evidence." *European Economic Review* 42 (6): 1033–67.

Cœuré, B. 2017. "Central Bank Communication in a Low Interest Rate Environment." *Open Economies Review* 28 (5): 813–22.

Cúrdia, V., A. Ferrero, G. C. Ng, and A. Tambalotti. 2015. "Has U.S. Monetary Policy Tracked the Efficient Interest Rate?" *Journal of Monetary Economics* 70 (March): 72–83.

de Mendonca, H., and G. de Guimarães e Souza. 2009. "Inflation Targeting Credibility and Reputation: The Consequences for the Interest Rate." *Economic Modelling* 26 (6): 1228–38.

Dovern, J., and G. Kenny. 2017. "The Long-Term Distribution of Expected Inflation in the Euro Area: What Has Changed Since the Great Recession?" ECB Working Paper No. 1999.

Ehrmann, M. 2015. "Targeting Inflation from Below: How Do Inflation Expectations Behave?" *International Journal of Central Banking* 11 (Supplement 1): 213–50.

European Central Bank. 2006. "Monthly Bulletin." (March).

———. 2016. "A Guide to the Eurosystem/ECB Staff Macroeconomic Projection Exercises." (July).

———. 2019. "The ECB Survey of Professional Forecasters — Third Quarter of 2019." (July).

Fischer, B., M. Lenza, H. Pill, and L. Reichlin. 2009. "Monetary Analysis and Monetary Policy in the Euro Area 1999-2006." *Journal of International Money and Finance* 28 (7): 1138–64.

Fujiwara, I. 2005. "Is the Central Bank's Publication of Economic Forecasts Influential?" *Economic Letters* 89 (3): 255–61.

Gerlach, S. 2007. "Interest Rate Setting by the ECB, 1999-2006: Words and Deeds." *International Journal of Central Banking* 3 (3, September): 1–45.

Gerlach, S., and J. Lewis. 2014. "ECB Reaction Functions and the Crisis of 2008." *International Journal of Central Banking* 10 (1, March): 137–57.

Gorter, J., J. Jacobs, and J. de Haan. 2008. "Taylor Rules for the ECB Using Expectations Data." *Scandinavian Journal of Economics* 110 (3): 473–88.

Hartmann, P., and F. Smets. 2018. "The First Twenty Years of the European Central Bank: Monetary Policy." ECB Working Paper No. 2219.

Hebden, J., and D. López-Salido. 2018. "From Taylor's Rule to Bernanke's Temporary Price Level Targeting." Finance and Economics Discussion Series No. 2018-051, Board of Governors of the Federal Reserve System. https://doi.org/10.17016/FEDS.2018.051.

Honkapohja, S., and K. Mitra. 2018. "Price Level Targeting with Evolving Credibility." CEPR Discussion Paper No. 12739.

Hubert, P. 2015. "Do Central Bank Forecasts Influence Private Agents? Forecasting Performance versus Signals." *Journal of Money, Credit and Banking* 47 (4): 771–89.

Hutchinson, J., and F. Smets. 2017. "Monetary Policy in Uncertain Times: ECB Monetary Policy since June 2014." *The Manchester School* 85 (S2): e1–e15.

Ikeda, T. 2010. "Time-Varying Asymmetries in Central Bank Preferences: The Case of the ECB." *Journal of Macroeconomics* 32 (4): 1054–66.

King, R., and A. Wolman. 1999. "What Should the Monetary Authority Do When Prices Are Sticky?" In *Monetary Policy Rules*, ed. J. B. Taylor, 349–404 (chapter 8). University of Chicago Press.

Kortela, T. 2016. "A Shadow Rate Model with Time-Varying Lower Bound of Interest Rates." Research Discussion Paper No. 19/2016, Bank of Finland.

Lyziak, T., and M. Paloviita. 2017. "Anchoring of Inflation Expectations in the Euro Area: Recent Evidence Based on Survey Data." *European Journal of Political Economy* 46 (January): 52–73.

———. 2018. "On the Formation of Inflation Expectations in Turbulent Times: The Case of the Euro Area." *Economic Modelling* 72 (June): 132–9.

Marcellino, M., and A. Musso. 2011. "The Reliability of Real-Time Estimates of the Euro Area Output Gap." *Economic Modelling* 28 (4): 1842–56.

Miles, D., U. Panizza, R. Reis, and A. Ubide. 2017. "And Yet It Moves. Inflation and the Great Recession." Geneva Reports on the World Economy, International Center for Monetary and Banking Studies (ICMB).

Neuenkirch, M., and P. Tillmann. 2014. "Inflation Targeting, Credibility, and Non-linear Taylor Rules." *Journal of International Money and Finance* 41 (March): 30–45.

Newey, W., and K. West. 1987. "A Simple, Positive Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55: 703–8.

Orphanides, A. 2003. "Historical Monetary Policy Analysis and the Taylor Rule." *Journal of Monetary Economics* 50 (5): 983–1022.

———. 2008. "Taylor Rules." In *The New Palgrave Dictionary of Economics*, 2nd edition, ed. S. Durlauf and L. E. Blume. London: Palgrave Macmillan.

Orphanides, A., and S. van Norden. 2002. "The Unreliability of Output-Gap Estimates in Real Time." *Review of Economics and Statistics* 84 (4): 569–83.

Sturm, J.-E., and J. de Haan. 2011. "Does Central Bank Communication Really Lead to Better Forecasts of Policy Decisions? New Evidence Based on a Taylor Rule Model for the ECB." *Review of World Economics* 147 (1): 41–58.

Surico, P. 2003. "Asymmetric Reaction Functions for the Euro Area." *Oxford Review of Economic Policy* 19 (1): 44–57.

———. 2007. "The Monetary Policy of the European Central Bank." *Scandinavian Journal of Economics* 109 (1): 115–35.

Svensson, L. E. O., and M. Woodford. 2005. "Implementing Optimal Policy through Inflation-Forecast Targeting." In *The Inflation Targeting Debate*, ed. B. S. Bernanke and M. Woodford (chapter 2). Chicago: University of Chicago Press.

Sveriges Riksbank. 2017. "Evaluation of the Riksbank's Forecasts." Riksbank Studies (May).

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

———. 2018. "Rules Versus Discretion: Assessing the Debate Over the Conduct of Monetary Policy." Economics Working Paper No. 18102, Hoover Institution, Stanford University.

van der Cruijsen, C., D.-J. Jansen, and J. de Haan. 2015. "How Much Does the Public Know about the ECB's Monetary Policy? Evidence from a Survey of Dutch Households." *International Journal of Central Banking* 11 (4, December): 169–218.

Wicksell, K. 1898. *Interest and Prices* (translated to English in 1936 by R. F. Kahn). London: Macmillan.

Williams, J. C. 2017. "Preparing for the Next Storm: Reassessing Frameworks and Strategies in a Low R-star World." FRBSF Economic Letter No. 2017-13, Federal Reserve Bank of San Francisco.

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

———. 2007. "The Case for Forecast Targeting as a Monetary Policy Strategy." *Journal of Economic Perspectives* 21 (4): 3–24.

Wu, 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.

Deposit Insurance and Banks' Deposit Rates: Evidence from the 2009 EU Policy*

Matteo Gatti[a] and Tommaso Oliviero[b]
[a]European University Institute
[b]University of Naples Federico II and CSEF

Deposit insurance is one of the main pillars of banking regulation meant to safeguard financial stability. In early 2009, the EU increased the minimum deposit insurance limit from €20,000 to €100,000 per bank account with the goal of achieving greater stability in the financial markets. Italy had already set a limit of €103,291 in 1994. We evaluate the impact of the new directive on the banks' average interest rate on customer deposits by comparing banks in the euro-zone countries to those in Italy, before and after the policy change. The comparability between the two groups of banks is improved by means of a propensity score matching. We find that the increase in the deposit insurance limit led to a significant decrease in the cost of funding per unit of customer deposit and that the effect is stronger for riskier banks, suggesting that the policy reduced the risk premium demanded by depositors.

JEL Codes: G21, G28.

1. Introduction

Deposit insurance is a widely employed measure intended to protect depositors, in part or in full, against the risk of insolvency. In the aftermath of the financial crisis and in order to protect

*We thank Domenico Curcio, Leonardo Gambacorta, Niels Johannesen, Agnese Leonello, Francesco Molteni, Steven Ongena, Lorenzo Pandolfi, Evi Pappa, Annalisa Scognamiglio, Saverio Simonelli, Davide Tomio, Alberto Zazzaro, and seminar participants at the EUI May Forum (May 2017), the SIE Annual Meeting (October 2017), the FEBS 2018 Conference (June 2018), and the first CefES International Conference on European Studies (June 2019) for helpful comments on various drafts of this paper. All remaining errors are our own. Corresponding author e-mail: tommaso.oliviero@unina.it.

the stability of their national banking systems, governments either adopted a deposit insurance scheme or redesigned the existing one by extending its scope and coverage (Demirgüç-Kunt, Kane, and Laeven 2015). Thus, the World Bank reported an increase in the number of countries that have some form of explicit deposit insurance scheme from 93 in 2013 to more than 107 in 2016 (Anginer and Bertay 2019). For example, the U.S. government temporarily increased its deposit insurance coverage from $100,000 to $250,000 in October 2008 and then made the change permanent in July 2010. Similarly, the European Union raised its deposit insurance limit from €20,000 to €100,000 in March 2009 as a first step toward the creation of a common regulatory environment within the European Banking Union.

The increases in deposit insurance limits are consistent with the theoretical foundation of an explicit deposit insurance as a bound on financial instability and, in particular, as a measure to avoid bank runs (Diamond and Dybvig 1983). In view of the fact that banks are mainly financed by deposits—they constitute about 60 percent of total bank assets in Europe and the United States—and that they play a vital role in the real economy, increasing deposit safety was crucial in reducing liquidity risk in distressed financial markets.

Beyond the short-run impact on financial stability, regulators have sought a more comprehensive understanding of the potential benefits and costs associated with increasing deposit insurance limits. The increase in the coverage limit affects not only financial stability but also the demand and supply of deposits. An increase in coverage may therefore affect the equilibrium interest rate on deposits and consequently banks' cost of funding (Cooper and Ross 2002; Davila and Goldstein 2015). While there is a long-standing debate in the literature regarding the impact of deposit insurance on the banks' cost of deposit funding, there are few attempts in the related empirical literature that use a quasi-experimental setting to examine the issue. This paper attempts to fill the gap.

We empirically evaluate the effect of the aforementioned March 2009 raising of EU deposit insurance coverage on banks' average deposit interest rates. In particular, we exploit a unique feature of the European change in policy, namely that Italian banks were already subject to a national deposit insurance limit of €103,291 which had been in place since 1994, well before the new European regime. Italian banks were therefore unaffected by the new EU policy. This allows us to employ a difference-in-differences (diff-in-diff) strategy by comparing Italian banks to those operating in the rest of the euro zone. We argue that the empirical results offer a reliable estimate of the causal effect of the increase in deposit insurance on banks' average deposit rates and deposit amounts. We find that the increase in the deposit insurance limit led to a decrease in the cost of funding per unit of deposits for euro-zone banks relative to banks in Italy of between 0.3 and 0.7 percentage point. This effect is sizable in terms of its macroeconomic implications. Indeed, Gambacorta and Shin (2018) recently showed that a lower cost of debt financing for banks implies a lower cost of equity and ultimately translates into large increases in bank lending.

We consider bank-level data provided by Bankscope (Bureau Van Dijk). In the treatment group, we consider banks operating in the European countries that jointly satisfy the following criteria: (i) they were part of the euro zone since the introduction of the euro as the single currency in January 2002; (ii) they were subject to an increase in deposit insurance limits between September 30, 2008 and June 30, 2009, and no further increase subsequently. The rationale for criterion (i) is that, given time fixed effects in all the regressions, we choose to select all countries that are subject to the same monetary authority. The rationale for criterion (ii) relates to our decision to limit the analysis to the two years subsequent to the policy (2009 and 2010) in order to reduce the risk of confounding policy changes or other exogenous events at the country level in the post-treatment years. The banks that operate in Italy were chosen for the control group not only because they were unaffected by the 2009/14/EC directive, but also because, as documented in Laeven and Valencia (2010), the Italian government did not adopt any specific policy response to the 2008 financial crisis. Furthermore, up until the 2011 sovereign debt crisis,[1] the Italian government did not intervene with any kind of extraordinary measure in order to directly assist its banks, such as capital injections or asset guarantees.

[1]For all practical purposes, the 2011 sovereign debt crisis was unexpected at the end of 2010, which is clearly demonstrated by the yield on Italian government bonds. To avoid the confounding effects of the sovereign debt crisis in our setting, we decided not to continue our empirical investigation beyond 2010.

We first estimate the impact of an increase in the deposit insurance limit using a diff-in-diff model, in which observations are weighted by individual bank size (measured by total assets). We consider for the dependent variable both interest rate expenses divided by total deposits, as a measure of average deposit interest rate, and the logarithm of total customers' deposits. By weighting each observation by the bank's total assets and including time and country fixed effects, the results can be interpreted as the macroeconomic impact of the policy. We find that there is a decrease in banks' average deposit interest rates of about 0.3 percentage point in treated countries relative to Italy and a relative increase in the total amount of bank deposits.

To extend the findings, we also estimate an unweighted diff-in-diff regression in which the unit of observation is at the bank level. This strategy allows us to include bank fixed effects in the regression analysis together with time fixed effects. Given that the sample of banks in the treatment and control groups is heterogenous along various dimensions, we strove for a sample of banks operating in the treated countries that is directly comparable to banks operating in Italy. To this end, we use a propensity score matching to identify banks within the treated countries that are comparable in terms of ex ante observable characteristics to banks operating in Italy. The results from the diff-in-diff estimation performed on the matched sample confirm that the policy had a negative and sizable impact on the average interest rates paid by banks on customer deposits (about 0.7 percentage point).

To further investigate the economic mechanism behind this finding, we exploit heterogeneities at the bank level. In particular, we find that the results for deposit rates are primarily driven by the riskier financial institutions. The empirical results suggest that an increase in deposit insurance coverage negatively affects per-unit cost of deposit funding and has either a positive or zero impact on the amount of deposits. Given that equilibrium movements are the results of changes in both the demand and supply of deposits, we conclude that the dominant effect is an increase in the supply of deposits and a reduction in the demand for deposits among the riskier banks. This conclusion is consistent with the predictions of the related theoretical literature.

On the supply side, depositors increase their supply of deposits following an increase in the deposit insurance limit, since they are willing to accept a lower interest rate for a given amount of deposits. In the absence of a direct impact of deposit insurance on banks' risk-taking due to moral hazard incentives, the increase in the deposit insurance limit should result in a lower deposit risk premium demanded by depositors (Bartholdy, Boyle, and Stover 2003). On the demand side, banks may either reduce or increase the deposit interest rate for a given level of deposits, depending on the set of assumptions regarding their incentives to take on risk. We distinguish between two groups of predictions in the literature. In the first, and in the wake of the seminal paper by Diamond and Dybvig (1983), Cooper and Ross (2002) study optimal deposit insurance limits in a setting where banks provide insurance to depositors in the form of deposit contracts. Since liquidity shocks faced by depositors constitute private information, late depositors can mimic the earlier ones and trigger a bank run. Cooper and Ross (2002) find that as the deposit insurance limit increases, depositors invest less in monitoring banks. As a consequence, and in the presence of moral hazard, banks increase the riskiness of their investments and their demand for deposits. Moreover, Matutes and Vives (2000) develop a model of banking competition based on Diamond (1984)'s delegated monitoring model and find that an increase in deposit insurance limits leads to an increase in the elasticity of the supply of deposits, thus increasing competition among the banks in the deposit market, which leads banks to pay a higher deposit rate for a given amount of deposits. In the second group, and if ex ante deposit insurance premiums are taken into account, larger deposit insurance coverage should decrease the banks' returns since less funds can be profitably invested. Since premiums increase with the level of deposit insurance, banks can pass these costs through to depositors and offer lower deposit rates per unit of deposits. As highlighted by Cooper and Ross (2002), the overall effect of the increase in the deposit insurance limit on banks' demand for deposits crucially depends on how it affects banks' risk-taking. Using our selected matched sample, we are able to study how a bank's realized risk responded to the increase in the deposit insurance limit in 2009. We find that treated banks did not show a significant difference in the Z-score relative to control banks, although we observe an increase in the ratio of nonperforming loans to total assets after 2009. At the same time, we find an increase, though not significant, in bank equity capitalization. Taken together, these results suggest that there is no substantial change in individual bank risk in response to the increase in the deposit insurance limit. This result is consistent with the recent evidence provided by Anginer, Demirgüç-Kunt, and Zhu (2014), who find that the effect of deposit insurance on risk-taking is dominated by its stabilization effect in turbulent times.

In addition to contributing to the debate regarding the impact of deposit insurance on banks' cost of deposit funding due to the resulting movements in the demand and supply of deposits, the findings also complement the recent empirical literature that focuses on the impact of deposit insurance on financial stability. Demirgüç-Kunt and Detragiache (2002) show that countries with an explicit deposit insurance system are associated with greater likelihood of a financial crisis, particularly when bank interest rates have been deregulated. In contrast, Dewenter, Hess, and Brogaard (2018) show that even in the presence of relatively homogenous banks, increases in deposit insurance limits have large effects on bank risk-taking in countries with weaker legal institutions. Our focus on the euro zone naturally means that we are analyzing a set of countries with relatively strong institutions and as a consequence we expect there to be less effect on risk. More recently, within-country studies have found a positive impact of deposit insurance on banks' risk-taking (see Lambert, Noth, and Schüwer 2017 for the United States and Chernykh and Cole 2011 for Russia). Finally, the paper contributes to the recent empirical evidence that depositors do indeed react to changes in deposit insurance limits (Iyer et al. 2017) and to changes in deposit insurance credibility (Bonfim and Santos 2018; Peia and Vranceanu 2017).

The rest of the paper is organized as follows: section 2 provides a description of the empirical setting and of the data; section 3 reports the results of the cross-country level analysis; and section 4 reports the results of the bank-level analysis. Section 5 concludes.

2. Institutional Setting and Data

2.1 The 2009/14/EC Directive

Directive 94/19/EC, which was signed on May 30, 1994, instituted a harmonized minimum level of deposit insurance on all deposits in the EU. Article 7 of the directive required all member states to put in place a deposit guarantee scheme that would cover up to at least €20,000 per depositor in the case that an insured bank could not redeem his deposit.[2] The €20,000 limit introduced by the 94/19/EC directive represented a first step toward the integration of European financial regulation, while at the same time leaving member countries with some level of discretion. Thus member countries could and did impose higher deposit insurance limits as they saw fit.

In response to the instability in the wake of the financial crisis, the floor was then raised on March 11, 2009 to €100,000 by the 2009/14/EC directive, whose goal was to enhance deposit protection and prevent liquidity issues from exacerbating banking instability. At the same time, the new directive supported the harmonization of the deposit insurance limits across Europe, a step toward the creation of the upcoming European banking union and of a European common deposit insurance scheme.

The actual convergence toward the €100,000 limit started a couple of months before the implementation of the 2009/14/EC directive. During the fourth quarter of 2008 and the first quarter of 2011, most European governments had already raised the limit of their deposit insurance coverage to €100,000. This increase was not, however, uniform across countries, ranging from no increase in Italy to an increase of €100,000 in Luxembourg and differing in the timing of the increase.

The uniqueness of the European case lies in the heterogeneity of the increase in deposit insurance limits among otherwise similar European countries. Thus, euro-zone countries shared a common currency and integrated markets for goods and services, but not a common banking union. Since the deposit insurance limit was not raised in all European countries, and since not all European countries were part of the euro zone, we restricted the sample to a subset of the most homogeneous countries. We selected European countries that satisfied both of the following criteria: (i) membership in the euro area since the introduction of the euro as the single currency in January 2002; (ii) their deposit insurance limit was raised between September 30, 2008 and June 30, 2009, with no additional

[2] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31994L0019:EN:HTML.

Table 1. Deposit Insurance Limit by Country in September 2008

Country	DI Limit (€)
Austria	20,000
Belgium	20,000
France	70,000
Germany	20,000
Greece	20,000
Ireland	20,000
Italy	103,291
Netherlands	20,000
Portugal	25,000
Spain	20,000

Note: The source of data is Demirgüç-Kunt, Kane, and Laeven et al. (2015) and data available online from various national deposit insurers.

increase in subsequent years. This selection strategy made it possible to compare banks in countries that were ex ante very similar to one another before the increase in the deposit insurance limit. Table 1 shows those countries that satisfy both of the conditions and their deposit insurance limits in 2008.[3]

Data on deposit insurance limits and their evolution through time have been collected from different sources. The main source of this data was a database assembled by Demirgüç-Kunt, Kane, and Laeven (2015). However, the database did not provide the exact date of when the deposit insurance limit was changed, so we supplemented the information using data available from various national deposit insurers in order to obtain the exact date and amount of the change.

2.2 Data Sources

Data were obtained from Bankscope, a Bureau Van Dijk (BVD) data set that contains yearly bank-level information with a global

[3]We do not have information on the main dependent variable, namely interest expenses per unit of customer deposits, for banks operating in Finland and Luxembourg, although they were part of the euro zone since the introduction of the euro as a single currency. Banks in those countries are not part of the sample.

coverage. Bankscope obtains balance sheet information from Fitch Solutions, the primary distributor of Fitch Rating content, and it also provides summary reports, peer analysis, and aggregated data reports. We focus on bank-level data for all banks that were active between 2006 and 2010 in the EU countries listed in table 1. We downloaded annual rather than quarterly data, since far fewer banks report quarterly data and they are less representative of the banking system in the respective country. The sample was restricted to banks with a consolidation code of C1, U1, or C2 and does not include banks with a consolidation code of U2. Codes C1 and C2 identify banks that report a consolidated financial statement that includes its controlled subsidiaries or branches, either with no unconsolidated companion (C1) or with an unconsolidated companion (C2). The financial statements of banks with code U1 include subsidiaries or branches with no consolidated companion.[4]

We further restrict the sample by considering only banks whose core business is credit intermediation, i.e., raising funds through deposits and lending them out as credit. We proceed in two steps, by first restricting the sample to banks labeled either as bank holding companies, commercial banks, cooperative banks, or saving banks as the peer group classification and then screening manually those banks according to their primary business line. Financial institutions whose main focus is on other activities, such as credit factoring and leasing, were dropped.[5]

We complemented the bank-level information from Bankscope with some country-level data publicly available from the OECD

---

[4]This selection criterion was used in order to maximize the data coverage of the main dependent variable (the ratio of interest expenses on customer deposits to total deposits) in our final data set. Indeed, if we had initially selected financial statements with the code U2 rather than C2, we would have lost observations for all major banks operating in Austria, Belgium, Germany, and Ireland and some of the major banks in the other euro-zone countries. We checked the robustness of the baseline results to our selection criterion by repeating the main regression analyses after imputing to the final sample of banks the unconsolidated balance sheet items (U2) of total customer deposits, total assets, and average interest rates (when available) for banks with consolidation code C2. The results are available on request.

[5]The initial sample consists of 200 banks operating in Austria, 30 banks in Belgium, 160 in France, 657 in Germany, 10 in Greece, 9 in Ireland, 451 in Italy, 22 in the Netherlands, 21 in Portugal, and 100 in Spain.

database, such as gross domestic product (GDP) growth, unemployment rate, and ratio of public debt to GDP. Finally, we added information on the level of covered deposits as a proportion of those deposits that are eligible for deposit insurance. We refer to this ratio as the covered-to-eligible (CtE) ratio, which is defined as the ratio of covered deposits in a particular country to the total amount of eligible deposits: $CtE = \frac{\text{Covered Deposits}}{\text{Eligible Deposits}}$. The greater its value, the higher is the proportion of deposits covered by national deposit insurance. The CtE ratio provides information on the potential effect on a country's banking system of a change in the deposit insurance limit: a low CtE ratio means that the amount of uncovered deposits is high and that a small change in the deposit insurance limit could turn uncovered deposits into covered ones. In other words, the lower is the CtE ratio, the greater is the amount of deposits that becomes covered after the change. Since European banks are not obligated to provide this information, we were forced to use estimates from a European Commission report by Cariboni and Uboldi (2008) on a country-level basis. These estimates are based on a survey of all EU member states in 2005.

Table 2 presents country-level summary statistics on some key variables for the European countries (treatment), in which the deposit insurance limit was increased following the 2009/14/EC directive and for Italy (control).

The statistics are computed as the mean for the pre-treatment years, i.e., 2006–08. The control group's statistics include only Italy. Column 3 in table 2 displays the mean difference between the average values in columns 2 and 1. The average deposit rate is measured by the interest expenses per unit of customer deposit of each bank weighted by each bank's total assets within each country-year. The cost of deposit funding (per unit of deposit) is an implicit measure of the average interest rate on deposits.[6] The weighted average deposit rate in Italy is significantly lower than in the European countries,

[6] A similar approximation was used recently by Gambacorta and Shin (2018), who measure the average cost of funds for a sample of banks by using the ratio between interest expenses and total interest-bearing liabilities (source: Bankscope). In line with our definition, differences in the cost of funds are captured by the average rate of interest the bank pays on its deposits and other interest-bearing liabilities. Demirgüç-Kunt and Huizinga (2004) also use this variable as an implicit measure of bank interest rates.

Table 2. Summary Statistics: Country Level 2006–08

	(1) Treatment	(2) Control	(3) Difference
Weighted Average Deposit Interest Rates	3.9456	2.7555	−1.1901**
Growth Rate of Total Deposits	0.0882	0.0927	0.0045
Growth Rate of Real GDP	0.0214	0.0225	0.0011
Unemployment Rate	7.2540	6.5250	−0.7290
Public Debt to GDP Ratio	63.4481	102.6700	39.2219**
Covered to Eligible Ratio	0.5554	0.7337	0.1783***
Ratio of Deposits for Top 25 Institutions	0.9261	0.8396	−0.0865
Observations	27	3	

Notes: Column 1 shows average values for the EU countries listed in table 1 except for Italy, whose average values are shown in column 2. Column 3 shows the mean difference between columns 2 and 1. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively. Weighted average deposit rate is the interest expenses per unit of deposit of each bank weighted by the bank's total assets for each country-year; the growth rate of total deposits is the yearly log change in the sum of each bank's total deposits in each country-year; the growth rate of real GDP (source: OECD) is the yearly log change of each country's real GDP; the unemployment rate and the public debt to GDP ratio are from OECD sources; covered to eligible ratio is taken from Cariboni and Uboldi (2008); the ratio of deposits for top 25 institutions is the ratio of total deposits of the 25 largest banks—defined using total assets—in each country-year to the sum of total deposits of all the banks in each country-year.

perhaps reflecting ex ante differences in deposit insurance regime. The growth rate of total deposits is measured as the annual average growth rate in the total amount of bank deposits aggregated by country. While average deposit rates in Italy are lower than in other countries before the policy was implemented, there is no significant difference in the growth rate of total deposits. Similarly, there is no difference in the growth rate of GDP and in the average unemployment rate, suggesting that the countries in our sample were experiencing similar macroeconomic trends before the policy change. Italy's ratio of public debt to GDP is larger than the average of other euro-zone countries. For this reason, we restrict the analysis to the years 2006–10, before the 2011 sovereign debt crisis affected some European countries, including Italy. The CtE ratio is larger in Italy than in the other European countries, suggesting that the amount of deposits below the limit is larger in Italy than in the rest of Europe. This finding is consistent with the higher deposit insurance limit in Italy than in other countries, since more eligible deposits are indeed covered by insurance, suggesting that the 2009/14/EC directive indeed increased the insurance coverage in euro-zone countries. Finally, the banking sector in Italy is slightly more competitive than in the rest of the sample countries. Thus, the average ratio of deposits held by the top 25 institutions (ranked by average total assets during the period 2006–08) is about 9 percentage points lower in Italy, suggesting less concentration in the deposit market, although the difference is not statistically different from zero.

In the next section, we estimate the impact of the increase in deposit insurance on the average (weighted) deposit interest rate and total customer deposits at the country level in a diff-in-diff analysis. The country fixed effects make it possible to de-mean the dependent variables and to exploit the within-country variation over time. The time fixed effects, on the other hand, make it possible to control for time factors that affect all euro-zone countries.

3. Country-Level Analysis

In this section, we provide preliminary cross-country evidence based on bank-level data. Since treatment occurs at the country level, we first run a cross-country analysis, weighting bank-level data by a bank's total assets within each country and imposing country fixed effects. The assumption underlying the analysis is that Italian banks were already subject to higher deposit insurance in 2009, thus providing a valid counterfactual to the trends of the other European banks absent the treatment. To study the causal impact of the EU directive in a diff-in-diff setting, we estimate the following equation:

$$y_{ict} = \gamma_c + \lambda_t + \delta(T_c \cdot post_t) + \epsilon_{ict}, \tag{1}$$

where y_{ict} is the outcome variable of interest for bank i at time t in country c. We consider both average deposit rates and the log of total deposits for this variable. γ_c captures the country fixed effect and λ_t captures the time trends. T_c is a dummy variable that equals 0 if country c is Italy and 1 otherwise, while $post_t$ is a time dummy

Table 3. Cross-Country Regressions

	(1)	(2)	(3)	(4)
	Average Deposit Rates		**Log of Deposits**	
T · Post	−0.3651*	−0.7388***	0.1250**	−0.0145
	(0.1764)	(0.1852)	(0.0496)	(0.0137)
Year Dummies	√	√	√	√
Country Fixed Effects	√	√	√	√
Weighted Regression	√		√	
Observations	3,123	3,123	8,390	8,390

Notes: T is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at country level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

that equals 1 in the post-reform period (after 2008). δ is the causal effect of interest. Notice that the estimates in this section are from weighted regressions, in which the observation of a bank is weighted by its total assets. This allows us to interpret the empirical estimates as a country-level effect of the policy. The results are in fact equivalent to a regression of weighted average deposit rates and weighted log of deposits at the $\{c, t\}$ level. The results in column 1 of table 3 show that the increase in the deposit insurance limit is followed by a reduction in average deposit rates of approximately 0.36 percentage point in treated countries relative to Italy (significant at the 5.9 percent level).[7]

[7]Our empirical results refer to changes in average deposit interest rates that cannot be directly linked to deposit interest rates charged on insured versus uninsured deposit accounts. However, Cannas et al. (2014) provide evidence of a substantial increase in the CtE deposit ratio for EU countries in 2011 with respect to 2005. The average CtE deposit ratio reached a level of about 70 percent for the euro-zone countries in our sample in 2011 and remained stable for Italy over the same period. Interestingly, the average CtE ratio for euro-zone countries in 2011 converged to the level of Italy. In light of this aggregate figure, our identified changes in the average deposit interest rates for euro-zone countries after 2008 can be, at least partially, imputed to a change in the share of insured to uninsured deposits.

Column 3 shows that the increase in the deposit insurance limit is followed by a relative increase in total deposits of approximately 12.5 percent. The difference in the number of observations in the two columns reflects the fact that the average deposit rate is not observable for many of the banks in the sample.[8]

We also estimate the unweighted version of the regression in equation (1) for average deposit rates (column 2) and logged deposits (column 4). In the unweighted version, the reduction in average deposit rates is stronger in magnitude and more precisely estimated, while the impact on total deposits is negative though not statistically significant. The comparison of the estimation results between the weighted and nonweighted regressions suggests that the drop in average deposit rates is greater for smaller banks, while the increase in deposits is greater for larger banks.[9]

The reliability of a diff-in-diff analysis naturally hinges on the parallel trend assumption, namely that the relative trends in deposit rates between treatment and control groups would have remained unchanged if there had not been any policy change. We provide an indirect test of this assumption by estimating the following equation:

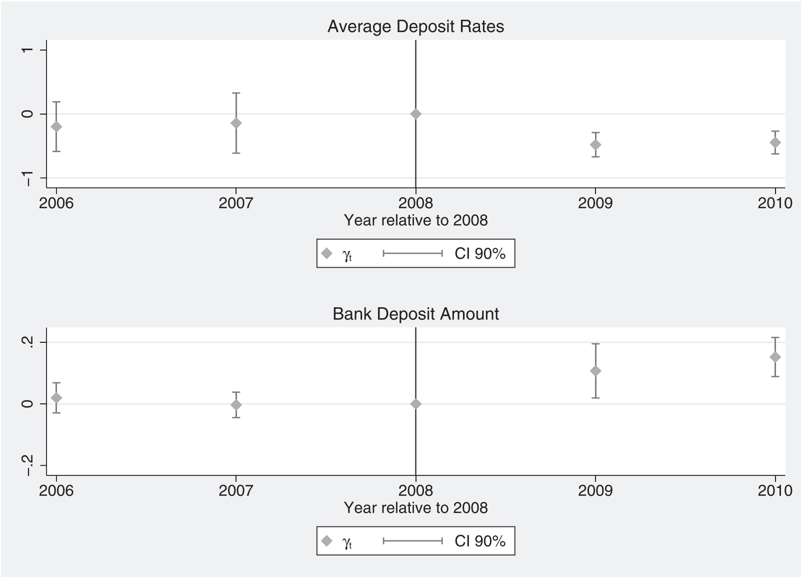
$$y_{ict} = \gamma_c + \lambda_t + \sum_{\tau=2006}^{2007} \gamma_\tau T_c 1(t=\tau) + \sum_{\tau=2009}^{2010} \gamma_\tau T_c 1(t=\tau) + \epsilon_{ict}, \tag{2}$$

where, unlike in equation (1), γ_τ are time-varying coefficients for the relationship between the outcome and T_c normalized relative to 2008, the year immediately prior to the policy change. Results are presented graphically in figure 1. Estimates for equation (2) are from a weighted regression with standard errors clustered at the country level.

[8] If we repeat the analysis employing only banks for which the deposit rate is observable, the results are practically unchanged.

[9] The results in table 3 are robust to the inclusion in equation (1) of additional control variables measured at country level, such as GDP growth rate, unemployment rate, and ratio of public debt to GDP. The robustness tests are not reported here, but are available upon request.

Figure 1. Event-Study Analysis: Country Level



Notes: The figure plots the pattern of the γ_τ coefficients from estimating equation (2) for average deposit rates as dependent variable (upper panel) and the log of deposits as dependent variable (lower panel). The capped lines show the 90 percent confidence interval on each coefficient relative to the reference year (2008).

Figure 1 presents two plots, each reporting the estimate from the above regression of the two different outcome variables: the average deposit rate (top panel) and the log of deposits (bottom panel). Each coefficient is normalized with respect to the estimated $\gamma_{\tau=0}$, where 0 is the value for 2008 and the reported confidence intervals are at the 90 percent level. The plots show that there is no statistically significant difference between the trends in either average deposit rates or total amount of deposits in 2006 and 2007 (pre-period). Consistent with the results displayed in table 3 and after controlling for country and time fixed effects, the average deposit interest rates decrease significantly in the treated countries relative to Italy after 2008, while the opposite occurs for total amount of deposits.

Taken together, the results indicate that the equilibrium reaction of average deposit rates (a price decrease) and of deposit amounts (a quantity increase) can be explained by a significant increase in supply of deposits following the increase in the deposit insurance limit. Since risk is being transferred away from depositors to the lender of last resort, deposits become more attractive relative to other investments, inducing depositors to increase their supply until the risk-adjusted return equals that of other asset classes (Sharpe 1994).

We consider the country-level analysis to be preliminary evidence of the macroeconomic changes caused by the increase in deposit insurance. To dig deeper into the identification of the effects of the EU policy change and to explain the economic mechanisms behind the empirical findings, we focus on a bank-level analysis in the following sections.

4. Bank-Level Analysis

In this section, we focus on bank-level data which provides a better understanding of the effects of an increase in deposit insurance limits on banks' average deposit rates and deposit amounts. In particular, we present results for the following empirical specification:

$$y_{ict} = \gamma_i + \lambda_t + \delta(T_c \cdot post_t) + \epsilon_{ict}, \tag{3}$$

where, unlike in equation (1), we include bank fixed effects to control for unobservable time-invariant bank characteristics (such as an individual bank's business model). Furthermore, we cluster standard errors at bank level and, unlike in the cross-country analysis, we give equal weight to each observation in the regression estimations. The estimation results for equation (3) are presented in table 4. The results in columns 1 and 2 are consistent with those for equation (1) in table 3.

However, these results must be interpreted with caution. Summary statistics reported in table 5 at the bank level for the treatment and control groups for the pre-period suggest strong unbalancing with respect to important variables that measure size, performance, and risk, including total assets, total deposits to total assets ratio,

Table 4. Bank-Level Analysis

	(1) Average Deposit Rates	(2) Log of Deposits
T · Post	−0.6404***	−0.0479***
	(0.1132)	(0.0161)
Year Dummies	√	√
Bank Fixed Effects	√	√
Observations	3,080	8,312

Notes: *T* is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at bank level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

Table 5. Summary Statistics: All Banks 2006–08

	(1) Treatment	(2) Control	(3) Difference
Total Assets (EUR Billions)	21.62	6.47	−15.15***
Average Deposit Interest Rates	3.47	1.74	−1.73***
Total Deposit to Total Assets Ratio	0.63	0.52	−0.11***
Tier 1 Capital to Total Assets Ratio	0.06	0.11	0.05***
Nonperforming Loans to Total Assets Ratio	0.02	0.04	0.02***
ROA	0.47	0.86	0.39***
Log(Z-score)	1.19	1.62	0.43***
Observations	3,651	1,370	

Notes: The table shows average values for banks operating in the EU countries listed in table 1 except for Italy (column 1) and banks operating in Italy (column 2); column 3 shows the mean difference between columns 2 and 1. Bank-level variables refer to key items from Bankscope. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

tier 1 capital to total assets ratio, nonperforming loans (NPLs) to total assets ratio, return on assets (ROA), and Z-score.[10]

[10]Following Laeven and Levine (2009), Z-score is calculated as the return on assets plus the capital-asset ratio divided by the standard deviation of asset returns in order to measure the distance from insolvency.

These pre-period differences in observable characteristics may potentially bias the empirical estimates of the coefficient of interest. In order to isolate the impact of the policy change, we seek a sample of banks that are ex ante comparable and that run on parallel trends in the pre-period. Therefore, in the next section, we build a sample of banks in the treated countries that is directly comparable to the sample of Italian banks by using a strategy based on the propensity score matching.

4.1 Propensity Score Matching

As previously mentioned, banks in the control group differ from treated banks in key financial variables. Therefore, a diff-in-diff analysis using Italian banks as a control group may lead to biased results. Differences in covariates between treated and control banks before the deposit insurance policy was implemented could affect the results since the control banks may not constitute a valid counterfactual for treated banks absent the treatment. In order to reduce this potential bias, we first run a propensity score matching that restricts the analysis to a more homogenous sample of treated and control banks and then use that subset in the empirical analysis.

In order to proceed with the matching strategy, we first average bank characteristics during the pre-treatment period (2006–08) since we wish to match treated banks with control banks in the years that preceded the policy change. We then restrict the control group according to the most relevant balance sheet variables. Finally, we match every bank in the control group with a bank in the treatment group that has the closest score, i.e., the same probability of being treated. We take their differences into account when estimating the following probit regression at the bank i − country c level:

$$M_{ic} = \alpha + \beta_1 Ta_{ic} + \beta_2 Tier1_{ic} + \beta_3 NPLs_{ic} + \beta_4 Dep_{ic} + \epsilon_{ic}, \qquad (4)$$

where M_{ic} is a dummy that equals 1 if bank i in country c is treated and 0 otherwise. Ta_{ic} is total assets, $Tier1_{ic}$ is tier 1 capital divided by total assets, $NPLs_{ic}$ is the ratio of nonperforming loans to total

Table 6. Summary Statistics: Matched Sample 2006–08

	(1) Treatment	(2) Control	(3) Difference
Total Assets (EUR Billions)	35.49	44.59	9.11
Average Deposit Interest Rates	4.31	1.95	−2.36***
Total Deposit to Total Assets Ratio	0.50	0.48	−0.02
Tier 1 Capital to Total Assets Ratio	0.08	0.09	0.01
Nonperforming Loans to Total Assets Ratio	0.02	0.02	−0.00
ROA	0.81	0.80	−0.01
Log(Z-score)	1.23	1.38	0.16
Observations	137	133	

Notes: The table shows average values for banks operating in the EU countries listed in table 1 except for Italy (column 1) and banks operating in Italy (column 2); column 3 shows the mean difference between columns 2 and 1. The number of banks in the treatment and control group after implementing the propensity score matching procedure is 48. Bank-level variables refer to key items from Bankscope. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

assets, and Dep_{ic} is the ratio of total deposits to total assets. Matching is then carried out using a nearest-neighbor approach with a caliper equal to 0.005. In other words, the propensity scores of two matched banks cannot differ by more than this value.[11] Finally, the matching is done without replacement, so that there is a unique match between a bank in the treatment group and a bank in the control group.

Table 6 shows the post-matching summary statistics of the treated group and the control group in the years 2006–08. The two matched samples each contain 48 banks. The two groups do not show any significant differences in size (measured by total assets), capitalization (measured by the ratio of tier 1 capital to total assets), business model (measured by the ratio of total deposits to total assets),

[11]In the benchmark procedure, the caliper threshold was chosen to minimize the dissimilarities between the control and treatment groups of banks and to maximize the sample size. In the appendix, we provide a discussion of this choice and provide a robustness exercise with a more conservative caliper of 0.001. The main results remain unchanged in this case.

and risk (measured by the ratio of NPLs to total assets). Although not directly used as regressors in the probit regression in equation (4), the two samples do not show any significant differences in performance (measured by ROA) or distance from insolvency (measured by the log of the Z-score). This result reinforces the comparability of the two groups in terms of ex ante observables. We finally observe that average deposit rates in the matched treated sample are greater on average than the matched control group. The difference of about 2.4 percentage points is consistent with a deposit insurance limit that is higher for the control banks than for the treated ones in the years 2006–08. This difference cannot, however, be completely attributed to different deposit insurance regimes. In fact, there may still be differences in unobservable characteristics at bank level that may explain the difference in the average cost of funding during the period 2006–08. In order to control for the potential impact of unobservable characteristics, we rely on time variation, which makes it possible to include bank fixed effects in the model. Therefore in what follows, we can rely on a diff-in-diff strategy that exploits the variation generated by the 2009 EU policy change.

Notice that the matching procedure was implemented without any restrictions on the initial sample of candidates to be matched. The final matched treatment and control samples are characterized by a significantly lower average of total assets than those of the largest European banks. It is in fact desirable to have fewer large banks in the sample since they may have been subject to direct government intervention in the aftermath of the global financial crisis. This situation enhances the internal validity of our strategy. As a further check, we compared the matched sample of banks to the list of banks provided by Laeven and Valencia (2010),[12] in order to verify that none of the banks had been subject to nationalization, recapitalization, or some other purchase or guarantee program implemented by a national government prior to 2009.[13]

[12] We refer to tables A.1 and A.3.

[13] The only exception is Aegon Bank N.V. located in the Netherlands, which had been recapitalized. We decided to keep this bank in the sample and performed a robustness check (not shown here, but available upon request) after which we excluded it from the sample.

Table 7. Bank-Level Analysis: Matched Sample

	(1) Average Deposit Rates	(2) Log of Deposits
T · Post	−0.7814**	0.0406
	(0.3420)	(0.0993)
Year Dummies	√	√
Bank Fixed Effects	√	√
Observations	334	448

Notes: The sample is obtained after implementing the propensity score matching procedure. *T* is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at bank level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

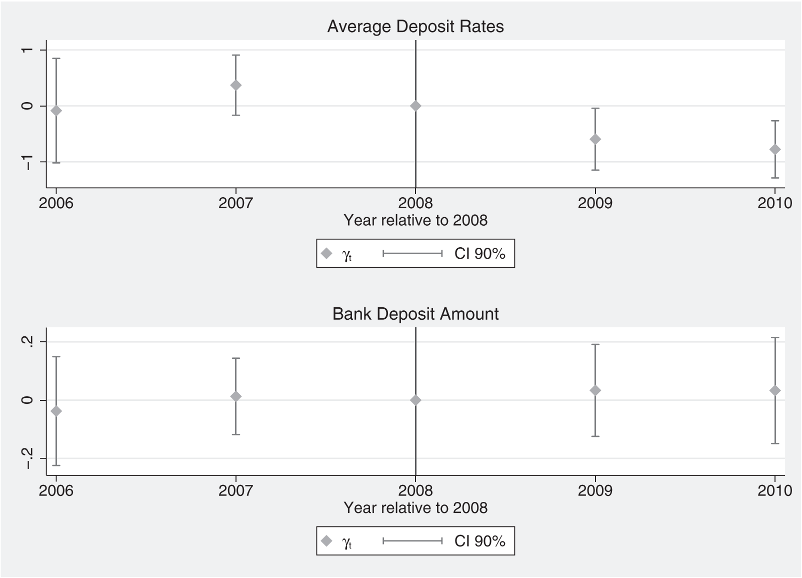
4.2 Empirical Results

Having obtained two matched samples of comparable banks, we can now estimate equation (3). Table 7 shows the estimation results for both average deposit rates and log of total bank deposits. Consistent with the results for the unmatched sample, we find a decrease in average deposit rates of approximately 78 basis points (column 1). In contrast to the previous results, the effect on total deposits is now positive but no longer statistically significant (column 2).[14]

We also test for the absence of pre-trends in the matched sample in order to assess our parallel trend assumption. Figure 2 indeed shows similar trends in the outcome variables before the policy change in 2008. The upper graph shows that the difference in average deposit rates between the treatment and control groups before treatment relative to 2008 is not significantly different from zero. In contrast, and consistent with the results presented in table 7, the

[14]As a robustness check, to control for the potential relationship between current and predetermined average deposit rates and amounts, we provide an additional test by augmenting the baseline model in equation (3) with the lagged dependent variable(s). That is, we estimate a version of the model in which we add the lagged values of average deposit rates and amounts as additional control variables. The results in table C.1 (in appendix C) show that the results are robust to this check since the coefficients of interest are only marginally affected in magnitude and significance.

Figure 2. Event-Study Analysis: Matched Sample



Notes: The figure plots the pattern of the γ_τ coefficients from estimating equation (2) for average deposit rates as dependent variable (upper panel) and the log of deposits as dependent variable (lower panel). The capped lines show the 90 percent confidence interval on each coefficient relative to the reference year (2008). The sample used in this analysis is after implementing the propensity score matching procedure.

difference decreases significantly in the years after the policy was introduced.

4.3 Impact on Risk

In order to clearly interpret the results, we also estimate the impact of deposit insurance on bank risk-taking. The empirical literature has discussed the effects of higher deposit insurance limits on financial stability. While Lambert, Noth, and Schüwer (2017) provide evidence that deposit insurance is associated with greater risk-taking, Anginer, Demirgüç-Kunt, and Zhu (2014) show that the higher risk induced by the moral hazard effect is dominated by the market stabilization effect in periods of greater financial distress. Given that

Table 8. Bank-Level Analysis: The Impact on Risk

	(1) Log(Z-score)	(2) $\frac{NPLs}{TA}$	(3) $\frac{Tier1}{TA}$
T · Post	−0.3146	0.0181*	0.0093
	(0.3400)	(0.0093)	(0.0079)
Year Dummies	√	√	√
Bank Fixed Effects	√	√	√
Observations	349	388	403

Notes: The regressions repeat the baseline analysis using the log Z-score (column 1), the nonperforming loans to total assets ratio (column 2), and the tier 1 to total assets ratio (column 3) as the dependent variable. The sample is obtained after implementing the propensity score matching procedure. T is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at bank level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

the 2009 EU policy change occurred in the aftermath of the global financial crisis, it is not ex ante clear which effect dominated in our setting.

In the following analysis, we show the effect of an increase in the deposit insurance limit on banks' realized risk after 2008. We estimate equation (3) using alternative measures of risk as dependent variables. Following Celerier, Kick, and Ongena (2018), we employ three measures of bank risk: the Z-score, which is inversely related to the probability of default (the greater the Z-score, the safer is the bank); the share of NPLs within total assets, which captures changes in the riskiness of the bank loan portfolio; and the ratio of tier 1 capital to total assets, which captures changes in the bank's degree of capitalization. Estimation results are shown in table 8.

We do not find a significant increase in the Z-scores for banks that operate in treated countries relative to the control banks as a consequence of the increase in the deposit insurance limit. At the same time, there is a significant increase in NPLs and a positive, though not significant, effect on the level of capitalization. We interpret this result as evidence that banks in treated countries experienced an increase in the risk of their loan portfolios; however, they responded by increasing their capitalization. The statistically insignificant effect on Z-scores suggests that the overall bank level risk did not change as a result of the increase in the deposit insurance limit.

4.4 Interpretation of the Results

The empirical findings indicate that the 2009 increase in the deposit insurance limit was followed by a significant decline in average deposit interest rates and by a not-significant change in total deposits and risk. The findings reveal that when the deposit insurance limit is increased, the average deposit interest rate decreases, that is, the banks' per-unit cost of deposit funding is reduced.[15] These findings can be explained by a joint shift in the supply and demand of deposits, as banks and depositors reacted to a change in the riskiness of the deposit contract.

The interpretation of the results is in line with the theoretical predictions. With respect to depositors, the results are consistent with an outward shift in the supply of deposits. An increase in deposit insurance limits moves risk away from depositors who now face lower risk in the event of default. Since depositors allocate their funds by comparing risk-adjusted returns in different asset classes (Sharpe 1994), deposits become more attractive as their risk-adjusted return increases. Depositors will thus increase their supply of deposits as long as the risk-adjusted returns on deposits are higher than returns in other asset classes. An increase in the supply of deposits lowers the return on deposits until the risk-adjusted return on deposits again equals that in other asset classes. Moreover, an increase in deposit insurance limits makes depositors less risk sensitive since they are more protected in the event of bank default. Since depositors are now less risk sensitive, they invest less effort in monitoring and are willing to increase their supply of deposits (Cooper and Ross 2002). In light of these arguments, and keeping

[15]The detected decline in average interest rate may potentially be correlated with changes in the maturity structure of deposits. Unfortunately, we cannot distinguish between the marginal interest rates offered on deposit contracts with different maturities, nor between interest rates offered on insured versus uninsured deposits. This type of analysis would require proprietary information on deposit accounts and interest rates charged on each account which is, unfortunately, not accessible.

demand constant, an outward shift in supply of deposits will result in lower deposit rates and an increase in total deposits.

With respect to the banks, the reaction to a higher deposit insurance limit can be twofold. On the one hand, a higher deposit insurance limit makes deposits a relatively more expensive source of funding. Banks are charged with a higher deposit insurance premium, which is needed to fund larger deposit insurance payouts in the event of a bank default. As deposits become more expensive, banks will only partially pass through higher costs to depositors, which will decrease the demand for deposits. On the other hand, an increase in deposit insurance limits can also encourage banks to increase deposit interest rates. Since depositors are now less risk sensitive and expend less effort to monitor banks, deposits become a less expensive source of funding for banks, which leads them to increase their demand for deposits (Cooper and Ross 2002). Moreover, higher deposit insurance limits may also raise the elasticity of supply of deposits, forcing banks to compete more actively for deposits by offering higher deposit rates (Matutes and Vives 1996).

The empirical results show a significant decrease in average deposit rates and a nonsignificant increase in total deposits. This is consistent with an increase in supply of deposits by depositors and a contemporaneous decrease in demand for deposits by banks. The increase in the deposit insurance limit may, in fact, have increased banks' operating costs, which—if not compensated for by higher risk-taking and profitability—would have reduced their demand for deposits.

4.5 Heterogeneous Policy Effects

The bank-level comparison of similar groups of treatment and control banks showed that an increase in the deposit insurance limit negatively affects the average interest rate on bank deposits, since depositors, who are now more protected against bank default, require a lower return for a given amount of deposits. If this is indeed the case, then we should observe a larger reduction in the average interest rate on deposits among riskier banks. In fact, risk-averse depositors allocate their money according to the risk-adjusted return offered by banks (Sharpe 1994) and require higher returns for riskier investments. As a result, riskier banks must offer higher rates to attract depositors and reward them for the greater risk they take on (Acharya and Mora 2015). An increase in the deposit insurance limit decreases the risk associated with deposits and makes them more homogenous across different levels of risk. Consequently, the decrease in average deposit rates after an increase in the deposit insurance limit should be larger among riskier banks until the risk-adjusted returns equalize.

In this section, we examine the heterogeneity between banks in order to understand whether the baseline effect on average deposit rates and deposit amounts is stronger for riskier banks. In particular, we consider a heterogeneity dimension which is likely to be viewed by depositors as a proxy for risk, namely the ratio of nonperforming loans to total assets.[16]

We test this intuition by ranking the banks according to percentage of NPLs within total assets and, in particular, we split the sample into four quartiles according to this measure of risk. We first estimate equation (3) while excluding the banks in the first quartile (i.e., the safest banks). We then exclude banks whose NPLs ratio is below the median value and, finally, we keep only the banks in the fourth quartile. The results when the dependent variable is the average deposit interest rate are reported in table 9. Column 1 reports the baseline results contained in table 7; columns 2–4 report the estimates for subsamples that feature progressively riskier banks.

According to the results, banks with a higher share of nonperforming loans within total assets show a larger decrease in average deposit rates. By comparing the magnitude of the coefficients across the different samples, it can be seen that, while the decline in average deposit interest rates in the baseline case is about 78 basis points, among the riskiest banks the decline is about 147 basis points, thus confirming that the drop in average deposit rates is larger for riskier banks.

We repeat the analysis with total deposit as the outcome variable. The results in table 10 show that the total amount of deposits does not significantly differ between the quartiles. In contrast to the baseline analysis, we observe a decline in total deposits among the riskiest banks, although the estimate is not statistically significant.

[16]In appendix B, we consider two alternative measures of perceived risk—the ratio of tier 1 capital to total assets and the log of the Z-score.

Table 9. Heterogeneity in Risk: Average Deposit Rates

	(1) Matched Sample	(2) >25	(3) >50	(4) >75
T · Post	−0.7814**	−0.7534**	−0.9462***	−1.4734**
	(0.3420)	(0.3554)	(0.3457)	(0.5247)
Year Dummies	√	√	√	√
Bank Fixed Effects	√	√	√	√
Observations	334	249	165	81

Notes: Column 1 presents estimates from the regression analysis reported in table 7; column 2 reports the estimates of a regression analysis using a subsample that excludes the banks whose $\frac{NPLs}{TA}$ is below the 25th percentile of the distribution; column 3 reports the estimates when excluding the banks whose $\frac{NPLs}{TA}$ is below the median; column 4 reports the estimates when excluding the banks whose $\frac{NPLs}{TA}$ is below the 75th percentile of the distribution. T is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at bank level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

Table 10. Heterogeneity in Risk: Total Deposits

	(1) Matched Sample	(2) >25	(3) >50	(4) >75
T · Post	0.0406	0.0278	0.0328	−0.0675
	(0.0993)	(0.0528)	(0.0742)	(0.1185)
Year Dummies	√	√	√	√
Bank Fixed Effects	√	√	√	√
Observations	448	332	220	110

Notes: Column 1 presents estimates from the regression analysis reported in table 7; column 2 reports the estimates of a regression analysis using a subsample that excludes the banks whose $\frac{NPLs}{TA}$ is below the 25th percentile of the distribution; column 3 reports the estimates when excluding the banks whose $\frac{NPLs}{TA}$ is below the median; column 4 reports the estimates when excluding the banks whose $\frac{NPLs}{TA}$ is below the 75th percentile of the distribution. T is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at bank level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

Taken together, the heterogeneity results confirm that the reduction in average deposit rates was stronger for banks that are perceived to be riskier by depositors, absent the deposit insurance coverage. Since lower average deposit rates are not associated with a higher amount of deposits, the results provide evidence of a larger reduction in demand for deposits among riskier banks.

5. Conclusions

This paper studies the effect of an increase in deposit insurance limits on average deposit rates by examining the effect of the 2009/14/EC directive, which increased deposit insurance limits in the European Union from €20,000 to €100,000. We exploit the fact that the limit had already been increased in Italy in 1994. According to the findings, the average deposit rates decreased substantially among banks in the EU relative to banks in Italy. The drop in average rates was larger among riskier banks, thus confirming the theoretical prediction that deposit insurance negatively affects banks' deposit interest rates by reducing the return required by depositors. Employing a combination of diff-in-diff and propensity score matching in order to provide a consistent estimate of the policy's impact, we conclude that the introduction of a higher limit on deposit insurance in the EU led to a significant reduction in the per-unit cost of deposit funding for euro-zone banks and to an alignment of banks' deposit rates across countries. Our findings contribute to the policy debate on the harmonization of the deposit insurance limits as a first step toward the creation of a common European deposit insurance scheme within the EU banking union.

Appendix A. Propensity Score Matching

In section 4.1 we carried out a propensity score matching that restricts the sample and allows us to compare banks with similar characteristics in the diff-in-diff estimation. The benefit of reducing the sample size to a group of more similar banks is to eliminate any potential bias that could be picked up by heterogeneous bank characteristics. On the other hand, it reduces the amount of information available and the precision of the estimates.

Table A.1. Summary Statistics: Matched Sample 2006–08, Robustness

	(1) Treatment	(2) Control	(3) Difference
Total Assets (EUR Billions)	32.68	44.16	11.48
Average Deposit Rates	4.30	1.83	-2.47^{***}
Total Deposit to Total Assets Ratio	0.49	0.50	0.01
Tier 1 Capital to Total Assets Ratio	0.09	0.09	−0.00
Nonperforming Loans to Total Assets Ratio	0.03	0.03	0.00
ROA	0.80	0.83	0.03
Log(Z-score)	1.12	1.50	0.38^{**}
Observations	76	76	

Notes: The table shows average values for banks operating in the EU countries listed in table 1 except for Italy (column 1) and banks operating in Italy (column 2); column 3 shows the mean-difference between columns 2 and 1. The matched sample is obtained using a propensity score matching procedure with a caliper of 0.001. The number of banks in the treatment and control group after implementing the propensity score matching procedure is 27. Bank-level variables are taken from Bankscope. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

In this appendix, we carry out an alternative and more conservative selection of the banks. This involves narrowing the caliper, i.e., the maximum distance in the score between a treated bank and a control bank, to 0.001. This should lead to a more homogeneous group of banks, although it reduces the size of the sample. Table A.1 provides the summary statistics for treated and control banks after the implementation of the propensity score matching with a caliper of 0.001. There does not appear to be a significant improvement in terms of the balance between the treatment and control groups with a smaller caliper, while there is a significant reduction in number of banks selected (from 96 to 54). Using this smaller sample, we then estimate the baseline diff-in-diff in equation (3) at the bank level.

Table A.2 presents similar estimates to the baseline regression although the standard errors are higher, possibly due to the smaller number of observations. This robustness shows that the choice of a caliper of 0.005 does not impose any costs in terms of balancing observable characteristics between treated and control banks and, at

Table A.2. Bank-Level Analysis: Matched Sample, Robustness

	(1) Average Deposit Rates	(2) Log of Deposits
T · Post	-0.9107^{*}	-0.1274
	(0.4766)	(0.1862)
Year Dummies	√	√
Bank Fixed Effects	√	√
Observations	184	250

Notes: The matched sample in the robustness check is obtained after implementing a propensity score matching procedure using a caliper equal to 0.001. *T* is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at bank level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

the same time, has the advantage of a larger sample size. In contrast, we also tried larger calipers when implementing the propensity score matching (for example, 0.01 and 0.1), which enlarged the matched sample of banks, but worsened ex post matching between control and treatment banks in terms of observable characteristics (results are not presented here for brevity but are available upon request).

Appendix B. Heterogenous Impact: Alternative Risk Measures

In this section, we repeat the heterogeneity analysis in section 4.5 using two alternative measures of risk: the ratio of tier 1 capital to total assets and the log(Z-score). We proceed in a similar manner by splitting the sample into four quartiles based on the distribution of the risk variables and then run the regressions on subsamples which progressively exclude the less risky banks. Column 1 in each table shows the results of the baseline regression, while the estimation results for the various subsamples appear in the subsequent columns. We sequentially exclude an additional quartile of the safer bank. Table B.1 shows the results when using average deposit interest rates as the dependent variable, while table B.2

Table B.1. Heterogeneity in Risk: Average Deposit Rates, Robustness

	(1) Matched Sample	(2) >25	(3) >50	(4) >75
T · Post	−0.7814**	−0.8611**	−0.3076	−1.2302*
	(0.3420)	(0.3655)	(0.3220)	(0.6681)
Year Dummies	√	√	√	√
Bank Fixed Effects	√	√	√	√
Observations	334	249	165	81

Notes: Column 1 reports estimates from the regression analysis reported in table 7; column 2 reports the estimates of a regression analysis using a subsample that excludes the banks whose $\frac{NPLs}{TA}$ is below the 25th percentile of the distribution; column 3 reports the estimates when excluding the banks whose $\frac{NPLs}{TA}$ is below the median; column 4 reports the estimates when excluding the banks whose $\frac{NPLs}{TA}$ is below the 75th percentile of the distribution. T is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at bank level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

Table B.2. Heterogeneity in Risk: Total Deposits, Robustness

	(1) Matched Sample	(2) >25	(3) >50	(4) >75
T · Post	0.0406	0.0247	−0.0044	0.2811
	(0.0993)	(0.1230)	(0.1997)	(0.4738)
Year Dummies	√	√	√	√
Bank Fixed Effects	√	√	√	√
Observations	448	336	220	112

Notes: Column 1 reports estimates from the regression analysis reported in table 7; column 2 reports the estimates of a regression analysis using a subsample that excludes the banks whose $\frac{NPLs}{TA}$ is below the 25th percentile of the distribution; column 3 reports the estimates when excluding the banks whose $\frac{NPLs}{TA}$ is below the median; column 4 reports the estimates when excluding the banks whose $\frac{NPLs}{TA}$ is below the 75th percentile of the distribution. T is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at bank level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

Table B.3. Heterogeneity in Risk: Average Deposit Rates, Robustness II

	(1) Matched Sample	(2) >25	(3) >50	(4) >75
T · Post	−0.7814**	−0.9153**	−0.4744	−0.8431
	(0.3420)	(0.4441)	(0.4384)	(0.6861)
Year Dummies	√	√	√	√
Bank Fixed Effects	√	√	√	√
Observations	334	250	168	83

Notes: Column 1 reports estimates from the regression analysis reported in table 7; column 2 reports the estimates of a regression analysis on a subsample which excludes the banks whose Log(Z-score) is below the 25th percentile of the distribution; column 3 reports estimates of a regression analysis on a subsample which excludes the banks whose Log(Z-score) is below the median; column 4 reports estimates of a regression analysis on a subsample which excludes the banks whose Log(Z-score) is below the 75th percentile of the distribution. *T* is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at bank level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

shows the result for total deposits. The results confirm that deposit insurance has a larger negative impact among banks in the fourth quartile, while there is no significant difference in the case of total deposits.

Similarly, tables B.3 and B.4 show results when a bank' riskiness is measured by the log(Z-score). With respect to average deposit rates, we find a larger negative effect when excluding the first quartile of the distribution (column 2), but not for banks above the median of the log(Z-score). In particular, the nonsignificant effect on average interest rates among the banks in the fourth quartile can be explained by the significant increase in total deposits observed in column 4 of table B.4. This result is theoretically compatible with an increase in the supply of deposits that is not matched by a sizable decrease in the demand for deposits by banks that are closer to default.

Table B.4. Heterogeneity in Risk: Total Deposit, Robustness II

	(1) Matched Sample	(2) >25	(3) >50	(4) >75
T · Post	0.0406	0.0858	0.1906	0.1413**
	(0.0993)	(0.1022)	(0.1255)	(0.0604)
Year Dummies	√	√	√	√
Bank Fixed Effects	√	√	√	√
Observations	448	334	221	110

Notes: Column 1 reports estimates from the regression analysis reported in table 7; column 2 reports the estimates of a regression analysis using a subsample that excludes the banks whose $\frac{NPLs}{TA}$ is below the 25th percentile of the distribution; column 3 reports the estimates when excluding the banks whose $\frac{NPLs}{TA}$ is below the median; column 4 reports the estimates when excluding the banks whose $\frac{NPLs}{TA}$ is below the 75th percentile of the distribution. T is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at bank level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

Appendix C. Cross-Border Banks

We check how many banks in our matched sample operate cross-border. As highlighted by Azevedo and Bonfim (2019), for cross-border banks that operate as subsidiaries in the European Union, the host deposit insurance fund is responsible for reimbursing insured depositors. We identify nine banks (six in the treated group and three in the control group) in our matched sample that are mother banks of cross-border subsidiaries. As a robustness check, we repeat the bank-level analysis after excluding those banks from the matched sample. For these banks, indeed, the consolidated balance sheet items that measure the cost of deposit funding could be potentially affected by changes in deposit insurance schemes in both the origin and the host jurisdictions. Results, reported in table C.2, are robust to this check. In particular, the effect of the increase in deposit insurance limit on average deposit rates is strongly significant, while no significant effect is detected on the total amount of customer deposits.

Table C.1. Bank-Level Analysis: Matched Sample, Robustness

	(1)	(2)	(3)	(4)
	Average Deposit Rates		Log of Deposits	
T · Post	-0.6796^{**}	-0.6587^{**}	0.0864	0.0178
	(0.3014)	(0.3031)	(0.0627)	(0.0516)
Lagged Average Deposit Rates	0.2411^{***}	0.0878		
	(0.0445)	(0.0651)		
T · Lagged Average Deposit Rates		0.1886^{***}		
		(0.0667)		
Lagged Log of Deposits			0.3966^{***}	0.1232
			(0.0946)	(0.1228)
T · Lagged Log of Deposits				0.3722^{***}
				(0.1281)
Year Dummies	√	√	√	√
Bank Fixed Effects	√	√	√	√
Observations	246	246	349	349

Notes: The estimates are based on a sample obtained using propensity score matching. *T* is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at bank level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

Table C.2. Bank-Level Analysis: Matched Sample, Robustness II

	(1) Average Deposit Rates	(2) Log of Deposits
T · Post	-1.0621^{***}	0.0710
	(0.3989)	(0.1031)
Year Dummies	√	√
Bank Fixed Effects	√	√
Observations	295	403

Notes: The sample is obtained after implementing the propensity score matching procedure and after excluding banks that operate cross-border. *T* is a dummy variable that equals 0 if the bank operates in Italy and 1 otherwise; *Post* is a time dummy that equals 1 in the post-reform years (2009–10) and 0 in the pre-reform years (2006–08). Standard errors appear in parentheses and are clustered at bank level. *, **, and *** denote significance at 10 percent, 5 percent, and 1 percent, respectively.

References

Acharya, V. V., and N. Mora. 2015. "A Crisis of Banks as Liquidity Providers." *Journal of Finance* 70 (1): 1–43.

Anginer, D., and A. C. Bertay. 2019. "Deposit Insurance." *ifo DICE Report* 17 (1): 3–8.

Anginer, D., A. Demirgüç-Kunt, and M. Zhu. 2014. "How Does Deposit Insurance Affect Bank Risk? Evidence from the Recent Crisis." *Journal of Banking and Finance* 48 (November): 312–21.

Azevedo, J. V., and D. Bonfim. 2019. "Deposit Insurance and Cross-Border Banks." *ifo DICE Report* 17 (1): 14–20.

Bartholdy, J., G. W. Boyle, and R. D. Stover. 2003. "Deposit Insurance and the Risk Premium in Bank Deposit Rates." *Journal of Banking and Finance* 27 (4): 699–717.

Bonfim, D., and J. A. C. Santos. 2018. "The Importance of Deposit Insurance Credibility." Working Paper.

Cannas, G., J. Cariboni, J., L. K. Veisari, and A. Pagano. 2014. "Updated Estimates of EU Eligible and Covered Deposits." Technical Report No. JRC87531, Joint Research Centre, European Commission.

Cariboni, J., and A. Uboldi. 2008. "Investigating the Efficiency of EU Deposit Guarantee Schemes." Technical Report No. JRC48944, Joint Research Centre, European Commission.

Celerier, C., T. K. Kick, and S. Ongena. 2018. "Taxing Bank Leverage: The Effects on Bank Capital Structure, Credit Supply and Risk-Taking." Working Paper.

Chernykh, L., and R. A. Cole. 2011. "Does Deposit Insurance Improve Financial Intermediation? Evidence from the Russian Experiment." *Journal of Banking and Finance* 35 (2): 388–402.

Cooper, R., and T. Ross. 2002. "Bank Runs: Deposit Insurance and Capital Requirements." *International Economic Review* 43 (1): 55–72.

Davila, E., and I. Goldstein. 2015. "Optimal Deposit Insurance." Working Paper.

Demirgüç-Kunt, A., and E. Detragiache. 2002. "Does Deposit Insurance Increase Banking System Stability? An Empirical Investigation." *Journal of Monetary Economics* 49 (7): 1373–1406.

Demirgüç-Kunt, A., and H. Huizinga. 2004. "Market Discipline and Deposit Insurance." *Journal of Monetary Economics* 51 (2): 375–99.

Demirgüç-Kunt, A., E. Kane, and L. Laeven. 2015. "Deposit Insurance around the World: A Comprehensive Analysis and Database." *Journal of Financial Stability* 20 (October): 155–83.

Dewenter, K. L., A. C. Hess, and J. Brogaard. 2018. "Institutions and Deposit Insurance: Empirical Evidence." *Journal of Financial Services Research* 54 (3): 269–92.

Diamond, D. W. 1984. "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies* 51 (3): 393–414.

Diamond, D. W., and P. H. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (3): 401–19.

Gambacorta, L., and H. S. Shin. 2018. "Why Bank Capital Matters for Monetary Policy." *Journal of Financial Intermediation* 35 (Part B): 17–29.

Iyer, R., T. Jensen, N. Johannesen, and A. Sheridan. 2017. "The Run for Safety: Financial Fragility and Deposit Insurance." Working Paper.

Laeven, L., and R. Levine. 2009. "Bank Governance, Regulation and Risk Taking." *Journal of Financial Economics* 93 (2): 259–75.

Laeven, L., and F. Valencia. 2010. "Resolution of Banking Crises: The Good, the Bad, and the Ugly." IMF Working Paper No. 10/146.

Lambert, C., F. Noth, and U. Schüwer. 2017. "How Do Insured Deposits Affect Bank Risk? Evidence from the 2008 Emergency Economic Stabilization Act. *Journal of Financial Intermediation* 29 (January): 81–102.

Matutes, C., and X. Vives. 1996. "Competition for Deposits, Fragility, and Insurance." *Journal of Financial Intermediation* 5 (2): 184–216.

———. 2000. "Imperfect Competition, Risk Taking, and Regulation in Banking." *European Economic Review* 44 (1): 1–34.

Peia, O., and R. Vranceanu. 2017. "Experimental Evidence on Bank Runs under Partial Deposit Insurance." ESSEC Working Paper No.1705.

Sharpe, W. 1994. "The Sharpe Ratio." *Journal of Portfolio Management* 21 (1): 49–58.

“Unconventional” Monetary Policy as Conventional Monetary Policy: A Perspective from the United States in the 1920s*

Mark Carlson and Burcu Duygan-Bump
Federal Reserve Board

To implement monetary policy in the 1920s, the Federal Reserve utilized administered interest rates and conducted open market operations in both government securities and private money market securities, sometimes in fairly considerable amounts. We show how the Federal Reserve was able to effectively use these tools to influence conditions in money markets, even those in which it was not an active participant. Moreover, our results suggest that the transmission of monetary policy to money markets occurred not just through changing the supply of reserves but, importantly, through financial market arbitrage and the rebalancing of investor portfolios. The tools used in the 1920s by the Federal Reserve resemble the extraordinary monetary policy tools used by central banks recently and provide further evidence on their effectiveness even in ordinary times.

JEL Codes: E52, E58, N22.

1. Introduction

In recent years, the Federal Reserve (the Fed), like many other central banks, has introduced new tools to implement monetary policy,

*We thank Jim Clouse, Barry Eichengreen, Bill English, Kenneth Garbade, Refet Gürkaynak, Christopher Hanes, Elizabeth Klee, Christoffer Koch, Jonathan Rose, Judit Temesvary, David Wheelock, and seminar participants at Bilkent University, Federal Reserve Board, Economic History Association Annual Meetings, Federal Reserve Bank of Chicago, the NBER DAE Summer Institute, and Rutgers University for valuable comments. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Board of Governors of the Federal Reserve or its staff. Author e-mails: mark.a.carlson@frb.gov and burcu.duygan-bump@frb.gov.

including large-scale asset purchases and use of administered rates. Several of these tools were introduced as unconventional and temporary policy tools, but some have argued that the Fed may have to rely on them more frequently going forward. That might be the case, for example, if there has been a decline in the long-run neutral real rate of interest—that is, the inflation-adjusted short-term interest rate consistent with keeping output at its potential on average over time—as suggested by Clarida (2014), Holston, Laubach, and Williams (2016), and Kiley (2018). Indeed, scholars and policymakers have increasingly engaged in a broader debate about how central bank policy should be implemented, including whether the central bank should use administered rates or target market rates, and whether the size and composition of the balance sheet should be used as policy tools (Stein 2012, Goodfriend 2014, Reis 2016, and Yellen 2016). An important part of the debate about which tools might be part of the toolkit is whether these tools are effective, and a growing literature has sought to evaluate their effectiveness, particularly of asset purchases (Krishnamurthy et al. 2011; D'Amico et al. 2012; Gertler and Karadi 2013; Altavilla, Carboni, and Motto 2015; and Haldane et al. 2016). However, such assessments are challenging given the limited number of actions and the fact that several asset purchase programs were announced following the financial crisis when market responsiveness may have been different than in normal times.

This paper provides a historical perspective on the tools available to the Fed by reviewing the U.S. monetary policy toolkit and analyzing the transmission of monetary policy to private money markets in the 1920s. Some of the tools used during this period as part of normal policy operations have important similarities to the tools introduced recently, even though the monetary policy framework in the 1920s was rather different.[1] Understanding their effectiveness as conventional policy tools during the 1920s might then provide additional insight on the potential effectiveness of such tools today

[1] In the 1920s for instance, the Fed was on the gold standard and had a minimum gold reserve ratio. Further, the Fed did not have a formal requirement to seek stable prices and full employment (this "dual mandate" was only codified in the Federal Reserve Act in 1977), but, as described in the 1923 Annual Report, the Board did believe it was appropriate for the Fed to operate some type of countercyclical monetary policy.

for use in ordinary times, not just in crisis times. In particular, the Fed implemented policy by adjusting administered interest rates and by purchasing both private and government securities.[2] While the asset purchases of Treasuries used by the Fed in the 1920s were of a smaller scale than the recent ones, the operations during two easing cycles in the earlier period did more than triple the Fed's Treasury portfolio.

Moreover, the interactions of the Fed with respect to one of the principle money markets of the time, the bankers' acceptance market, offers useful insights into the channels by which monetary policy was transmitted to private financial markets. Traditionally, most work on monetary policy transmission in the 1920s has focused on the reserves channel—how the Fed's actions affected the supply and cost of reserves. We test the reserves channel, but we also investigate whether other channels that have been of interest in modern times, such as the portfolio balance channel often associated with large-scale asset purchases, may have mattered as well. In doing so, this paper also adds to the literature on understanding the channels through which monetary policy is transmitted to financial markets.[3]

We start with a review of the tools available to the Fed in the 1920s to implement monetary policy and a discussion of how they functioned.[4] The monetary policy toolkit at this time is particularly interesting because the Fed had three policy instruments at its disposal, each of which worked in a slightly different fashion. The first tool was the rate the Fed charged for discount window loans (or rediscounts) where banks could choose to borrow from the Fed

[2]For ease of exposition, we typically refer to the Fed as a single entity when it comes to setting monetary policy. The reader should keep in mind that this is a simplification. The regional Reserve Banks had considerable autonomy regarding open market operations, although there was a committee of Reserve Bank officials that tried to coordinate purchases and sales. The Reserve Banks had authority to set their discount rates independently and were not always in sync with each other regarding changes in these administered policy rates; there were also times when discount rate changes sought by the Reserve Banks were not approved by the Federal Reserve Board. In this paper, we focus on policy rates prevailing at the Federal Reserve Bank of New York and the systemwide holdings of securities.

[3]See, for instance, Hamilton 1997; Demiralp, Preslopsky, and Whitesell 2006; Carpenter, Demiralp, and Senyuz 2016; Duffie and Krishnamurthy 2016.

[4]As we discuss in more detail below, we use the term monetary policy to refer to actions taken by the Fed and the consequences of those actions. We touch only lightly on the intent behind such actions.

to obtain reserves. The second tool was the rate at which the Fed would purchase bankers' acceptances; this tool directly affected the value of a particular money market instrument and affected the cost (and incentive) for banks to obtain reserves by choosing to sell the acceptances to the Fed. The third tool was open market operations in government securities in which the Fed would add or subtract reserves through purchase or sale of Treasury securities at the going market rate.

After describing the three policy instruments and how they worked, we examine their effectiveness in influencing conditions in money markets during the 1920s. The connection between changes in the policy tools and changes in conditions in private money markets is a key link in the transmission of monetary policy. We focus on the years from 1923 until 1929—after distortions from war finance needs had diminished and before the onset of the financial distress of the Great Depression. In particular, we test whether private money market rates in New York City, where the major money markets in the United States were located, responded to changes in the discount window rate of the Federal Reserve Bank of New York (New York Fed), changes in the discount at which the New York Fed purchased acceptances, and changes in the System's holding of government bonds.

We find that the policy instruments were effective in influencing private money market rates, where higher administrative rates tended to raise private interest rates while purchases of Treasury securities reduced private interest rates. Indeed, we find that the impact of large-scale asset purchases on money markets was fairly substantial and larger than those that have been found for the recent asset purchase programs. Such results are in line with those of Bordo and Sinha (2016), who find that such purchases had substantial effects on Treasury yields, which they then estimate had notable macroeconomic effects. However, these results are in contrast to the arguments in Toma (1989), who documents that the impact of open market operations on Federal Reserve credit was always offset by changes in bank borrowing through the discount window. This was famously labeled "the scissors effect" by Friedman and Schwartz (1963) and was recognized by Reifler (1930) and Burgess (1936).

To gain additional insight into the channels through which monetary policy was operating and to understand why we find that monetary policy was more effective than what would be implied by this scissors effect, we focus on the bankers' acceptance market and the Fed's rate for purchasing these instruments. Specifically, we test whether the bankers' acceptance market was sufficiently developed such that it enabled the Fed to use changes in its acceptance rate to influence conditions in money markets through channels in addition to the reserves channel. Certainly changes in the rate at which the Fed purchased these securities affected the incentives of banks to sell the acceptances to the Fed and thus the availability of bank reserves and financial conditions. However, it is possible that there were additional effects. In the 1920s, the Fed supported the growth of the acceptance market with the intent that banks and other financial institutions use these instruments as part of their liquid investments. To the extent that other institutions adjusted holdings and prices of other instruments as the rates on acceptances changed, then arbitrage and portfolio rebalancing may have meant that changes in the rate at which the Fed bought these securities had sizable impacts on the prices of other money market securities. Balabanis (1935) suggests that such behavior occurred to some extent. Alternatively, it may be that the price of acceptances would not necessarily have mattered for any other market prices because the Fed was intervening strongly in this market, holding more than 40 percent of the outstanding amount of acceptances at times.

We test the channels of transmission by looking at whether the changes in reserves appear to account for most of the change in money market rates. While we find that the effect of purchases of Treasury securities and bankers' acceptances were of similar magnitudes, consistent with the reserves channel, we also show that, for acceptances, the change in the rate at which the Fed purchased these securities had much larger effects than could be explained by just the reserves channel. Instead our evidence is consistent with additional transmission channels, such as portfolio balancing, being important during the 1920s.

In addition, we find that banks' holdings of acceptances responded to changes in the acceptance rate and that the commercial paper rate responded to changes in outstanding amounts of acceptances. These findings are suggestive of a portfolio balance channel and are consistent with arbitrage between the acceptance market and the commercial paper market being a channel through which monetary policy operated. They also support the idea that the acceptance market, while strongly influenced by the Fed, was integrated with other money markets.

The paper is organized as follows. Section 2 describes the Fed's monetary policy toolkit, and discusses the implications of the use of these tools for the Fed's balance sheet and the channels through which they influenced money markets. In section 3, we provide estimates of the size of the effects of the different monetary policy tools on money market interest rates in the 1920s. Section 4 explores the potential transmission channels, and particularly whether the changes in the acceptance rate operated through a portfolio balance channel in addition to its impact on reserves. Section 5 concludes.

2. Monetary Policy Toolkit and Transmission

In this section, we review the three main tools the Fed used to implement monetary policy in the 1920s—the discount window, purchases of bankers' acceptances, and open market operations in government securities—and how they shaped the Fed's balance sheet.[5] We also discuss several mechanisms through which changes in these tools may have been transmitted to private money markets.

Note that in this paper, our focus is the Fed's toolkit and how the monetary actions taken by the Fed affected money market conditions. We are neither focused on the reasons for the policy actions nor whether the policymakers had the correct intentions behind their actions. Friedman and Schwartz (1963), Wicker (1966), Wheelock (1991), and Meltzer (2003) provide a detailed review of the factors driving monetary policy in this period.

It is worth noting that during this period, the Fed was required to maintain a legal gold reserve equal to 40 percent of notes in circulation plus 35 percent of deposit liabilities. If that requirement would have been binding, or close to binding, it would have implications for our analysis. However, the minimum gold ratio in the period we analyze was equal to 60 percent of circulation and deposits, well above the statutory requirement.

It is also worth remembering that the Fed in this period was generally operating a countercyclical monetary policy with a stated

[5] A more detailed description of the tools may be found in Carlson and Duygan-Bump (2016).

goal of accommodating domestic commerce and business, without allowing speculative excesses to create instability. In this sense, the Fed was doing more than simply offsetting the effects of gold and currency flows on bank reserves as would be implied by the gold standard (Friedman and Schwartz 1963). The Fed would tighten policy when they viewed credit growth as excessive, and ease when industry and trade were in need of support. Conditions in financial markets were viewed as signals about the demand and supply of credit growth. Wheelock (1991) notes these signals required careful interpretation; for instance, low money market rates could signal low loan demand as well as excessive supply of reserves, but the appropriate monetary policy response differs depending on the underlying reason. Consistent with Friedman and Schwartz (1963)'s view of the 1920s as the "high tide" of the Federal Reserve System, Wheelock also suggests that the Fed interpreted the signals correctly in the 1920s given the correlation between monetary policy and industrial production in this period.

2.1 The Discount Window

In the 1920s, one of the primary tools for implementing policy was the discount window, where the Fed could (re)discount paper for banks or provide advances (loans) against eligible collateral.[6] Indeed, member banks, as a group, appear to have needed to borrow regularly from the Fed in order to meet their reserve requirements. The rates that were charged for providing credit through the discount window could be increased or decreased in order to affect credit conditions.

The operations of the discount window were overseen by the 12 Federal Reserve Banks, and the rates that were charged at the window were set by these banks subject to approval by the Federal Reserve Board. As interbank markets were not as integrated in this

[6]Eligible collateral in the 1920s consisted of only government securities and short-term commercial, agricultural, or industrial paper that was used to produce, purchase, carry, or market goods (see Bordo and Wheelock 2013). The Fed could not discount private promissory notes, such as corporate bonds or longer-maturity commercial and industrial loans. Thus, there was a much narrower range of eligible collateral than is the case today. For details of rediscounting, see Hackley (1973).

period as they are today, the Federal Reserve Banks had some scope to set discount window rates that differed across districts.[7] As our analysis focuses on the money market rates in New York, we focus on the discount window rate at the New York Fed.[8]

The rates that the New York Fed charged on its discount window loans were often close to, but below, the interest rates on private money market rates in New York. These were markets in which banks were typically lenders. The discount window rate was often above the rates the money center banks in New York paid on their deposits, including their interbank deposits. As described by the New York Clearing House Association (1920), the maximum rates that member banks were allowed to pay on interbank deposits and certificates of deposit issued to banks in the early 1920s were both below and a function of the discount rate of the New York Fed.[9] The relative expensiveness of the discount window is consistent with Burgess's report that the New York banks would often repay their discount window loans quickly when they received additional funds due to gold flows or Fed asset purchases (Burgess 1936, pp. 235–36).

2.2 Open Market Purchases of Bankers' Acceptances

The Fed could purchase bankers' acceptances in the open market as part of its open market authority.[10] The primary use of bankers'

[7]Efforts by banks to arbitrage differences in discount window rates promoted the early development of the federal funds market and the subsequent integration of interbank markets (Turner 1931).

[8]The reserve banks could, and did, have multiple discount window rates that varied with the collateral being used. During World War I, all the reserve banks offered a preferential rate for loans backed by government securities to bolster demand for such securities. By the early 1920s, the New York Fed had a single discount window rate.

[9]Specifically, the rules stated that the maximum rate that member banks could pay on such deposits was set at 1 percent when the 90-day discount rate for commercial paper at the New York Fed was 2 percent. For each increase of one-half percentage point in the discount rate above 2 percent, the maximum rate that member banks could pay was increased by one-quarter of a percentage point (New York Clearing House Association 1920, p. 16). While the strict relationship between rates and the New York Fed's discount window rate was relaxed in the mid-1920s, it appears that rates the money center banks paid on many of their deposits remained below the discount window rate.

[10]While the Fed could purchase a slightly broader set of securities than just bankers' acceptances, acceptances constituted nearly all the paper bought by the

acceptances in the 1920s was as a money market instrument to finance trade, especially international trade (Beckhart 1932). When an exporter shipped goods abroad, they typically had to wait to be paid until the goods reached the market and were sold. Rather than wait, the exporter could have brought a bill indicating the shipment to his bank and received a loan against that bill. Banks financed such loans by endorsing the bills and bringing them to larger banks, usually in a money center. The money center banks would then "accept" the bill and provide money to the exporter's bank. The money center bank could hold that bill or sell it into the market as a bankers' acceptance. The acceptance was guaranteed by the payment the exporter expected to receive, the promise of the exporter's bank to make good on the paper if the exporter failed, and the promise of the money center bank to make good on the paper if the exporter's bank failed. This type of instrument was little used in the United States prior to the creation of the Fed. Indeed, banks with national charters were forbidden to issue such securities. As many prominent European money markets, such as London, had large bankers' acceptance markets and because that these securities backed "real transactions," the founders of the Fed were keen to develop this market in the United States, which would also help promote the U.S. dollar as an international currency (Ferderer 2003; Eichengreen and Flandreau 2012). These efforts were generally successful and by the mid-1920s, this market was comparable in size to other important money markets.[11]

The Fed had a passive role in the open market operations in bankers' acceptances. Instead of buying a certain amount of acceptances directly from the market, the Federal Reserve Banks would set

---

Fed, so we use the term bankers' acceptances to refer to all such paper. The Fed still has the authority to purchases bankers' acceptances; however, it has not used this authority in some time. Open markets operations in bankers' acceptances and the use of repurchase agreements on bankers' acceptances to manage reserves were ceased in 1977 and 1984, respectively. See Small and Clouse (2005).

[11] In 1926 the average monthly outstanding volumes of the different money market instruments were bankers' acceptances: $691 million; commercial paper: $627 million; brokers' loans on call: $2,288 million; and brokers' loans on time: $825 million. There was an emerging federal funds market, but that was very small. There was also a repo market involving a variety of collateral, including bankers' acceptances, but the size of this market is unknown. For additional details on these markets, see Beckhart (1932).

the rates at which they would buy acceptances of particular maturities, where this rate was the discount relative to the face value of the acceptance; we refer to this discount as the acceptance rate. The Federal Reserve Banks would then take all eligible acceptances that were delivered to them.[12] Given the desire of the Fed to promote this market, the rates at which the Federal Reserve Banks would buy acceptances tended to be set favorably relative to the rate at the discount window. As with rates at the discount window, the 12 reserve banks could each set their own rates for purchasing these securities. Again, given our focus on the New York markets, we use the rates offered by the New York Fed. (The acceptance market in New York was the largest market for these securities.)

The Fed's holdings of acceptances represented a significant share of the market; at times the Fed held nearly 40 percent of outstanding acceptances. Consistent with its large share of the market, the rate at which the New York Fed was willing to buy acceptances heavily influenced the market price. As shown in figure 1, the market rate in New York and the rate at which the New York Fed offered to purchase the securities were almost identical.

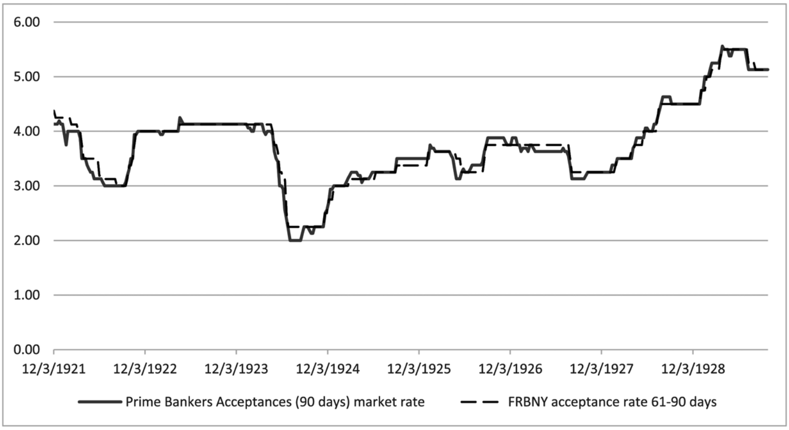
2.3 Open Market Operations in Government Securities

The Fed also has the authority to purchase or sell government obligations in the open market. Efforts to coordinate operations across the System eventually resulted in the creation of the Open Market Investment Committee in 1923.

Around this time, large-scale open market operations in government securities began to be seen as a tool that could be used to manage the aggregate quantity of credit and support the discount window policy. Some shifts in policy were associated with large swings in the holdings of government securities. For instance, in 1924, the Fed took steps to ease monetary policy and increased its holdings of Treasury securities from $100 million in January to

[12]It is not clear what the maturity structure of these instruments at origination was, but at particular points in time, roughly 40 percent of the holdings of the Fed had maturities of less than 15 days, 20 percent had a remaining maturity of between 16 and 30 days, and 25 percent had a remaining maturity of 31 to 60 days. The rest had a maturity of more than 60 days.

Figure 1. Market Rate of Interest and FRBNY Buying Rate on Bankers' Acceptances



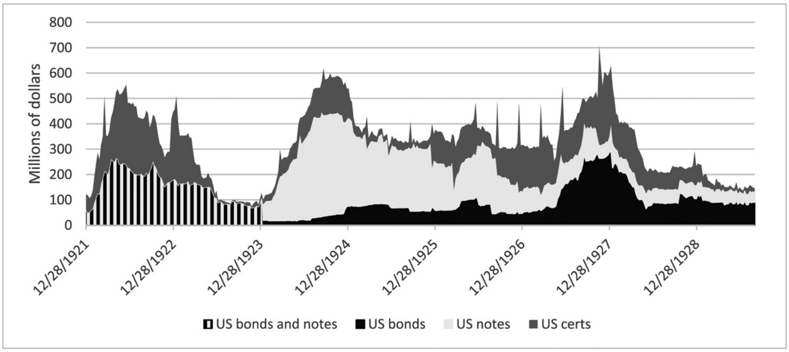
Source: Federal Reserve, Banking and Monetary Statistics, 1943.
Note: FRBNY is the Federal Reserve Bank of New York.

$600 million by October. Purchases were not announced but were observable from the Fed's weekly publication of its balance sheet. The Fed's holdings of Treasury securities consisted of the full range of Treasury securities: certificates of indebtedness (with maturities of a year or less), notes (with maturities of between 1 and 10 years), and bonds (with maturities of more than 10 years).[13] The maturity of the securities that were purchased or sold by the Fed when it engaged in substantial operations in Treasury securities appears to have varied over the 1920s, as shown in figure 2. (This is in contrast to the large-scale purchases by the Fed following the Great Recession which focused on purchases of longer-term securities.)

With the Open Market Investment Committee, most purchases were conducted by the New York Fed and allocated to the accounts of the other Reserve Banks.

[13]During this period, the Fed occasionally purchased cash-management securities directly from the Treasury. Such purchases were rare and short term, but they were large when they occurred. See Garbade (2014) for details.

Figure 2. Federal Reserve Holdings of Treasury Securities by Maturity



Source: Federal Reserve, Factors Affecting Bank Reserves and Condition Statement of Federal Reserve Banks (H.4.1 statistical release).

2.4 Balance Sheet of the Federal Reserve

The balance sheet of the Fed for year-end 1926 is shown in table 1. From this table, it is clear that the major asset of the Federal Reserve System was gold, which is consistent with the United States being on the gold standard. It is also apparent that, during this period, the direct exposure of the Fed to the condition of the commercial banking sector was fairly substantial. Private credit, consisting of bills of acceptance purchased in the open market, paper rediscounted, or advances to member banks on acceptable collateral (often government securities), constituted about 20 percent of the Fed's assets. Advances, which were mostly secured by U.S. government securities, were typically more substantial than discounts, which were typically of commercial and agricultural paper. Purchases of bankers' acceptances securities represented 7 percent of the Fed's assets. Most of the acceptances held by the Fed were associated with imports or exports, though some were associated with domestic trade or inventory finance. Holdings of government securities were a smaller share of the Fed's balance sheet in this period, averaging about seven percent of total assets.[14]

[14] At $315 million, the Fed's holdings of Treasury securities represented about 1.7 percent of outstanding interest-bearing Treasury securities.

Table 1. Balance Sheet of the Federal Reserve: December 31, 1926

Assets		Equity and Liabilities	
Gold Held against Federal Reserve Notes	1,448.6	Federal Reserve Notes	1,850.8
Other Gold	1,369.9		
Bills Rediscounted		Deposits	
Commercial and Industrial Paper	170.6	Member Bank Reserves	2,194.1
U.S. Government	1.1	Other Deposits	81.9
Other Bills	3.1		
Advances		Other Liabilities	669.7
Secured by U.S. Government	364.2		
Otherwise Secured	97.6		
Bills Bought in Open Market			
Acceptances for Imports and Exports	252.2		
Acceptances for Domestic Trade	77.7		
Other Bills	51.0		
U.S. Gov. Securities Bought in Open Market			
Certificates of Indebtedness	179.5		
Notes	87.4		
Bonds	48.0		
Uncollected Items	730.5	Capital	124.8
Other Assets	268.7	Surplus	228.8
Total	5,150.1	Total	5,150.1

Source: Federal Reserve Board Annual Report for 1926.
Note: All values in millions of dollars.

2.5 Channels by which Monetary Policy Was Transmitted to Financial Markets

In this section, we describe some of the channels through which monetary policy could have been transmitted to financial markets. While there are many channels through which monetary policy operates, such as the bank lending or the exchange rate channel, we focus especially on two channels: the reserves channel, which was emphasized by contemporaries, and the portfolio balance channel—which has been associated with the more recent large-scale asset purchases employed by several central banks. We also briefly touch on the signaling channel.

2.5.1 The Reserves Channel

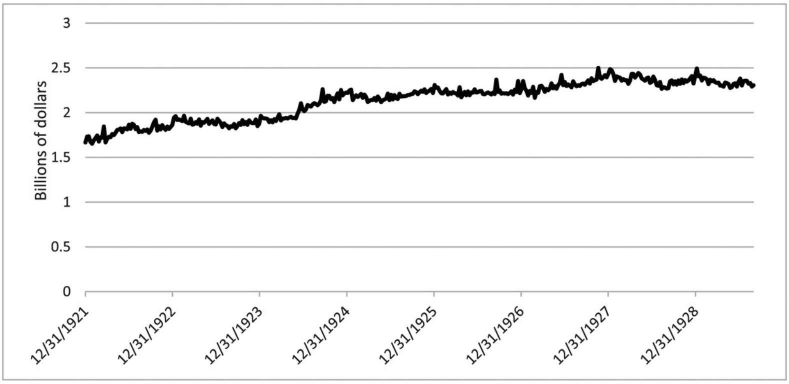
The transmission channel described by contemporaries (especially Reifler 1930 and Burgess 1936) is the reserves channel, where the Fed affected the money market by changing the supply and cost of obtaining reserves.[15] Banks demanded reserves in order to comply with reserve requirements and to facilitate their conduct of the business of banking; for instance, reserves are useful for processing payment system transactions. The supply of reserves was determined importantly by the Fed through its open markets operations. Purchases (or sales/maturing) of Treasury securities would expand (contract) the supply of reserves. Open markets operations in acceptances required that the banks decide to sell such securities to the Federal Reserve Bank, but the Federal Reserve Bank could influence commercial banks' incentives by changing the discount at which it purchased them. As these securities were short term and naturally "self-liquidating," they would mature over time and tighten policy unless they were replaced.[16] Commercial banks could also obtain reserves by borrowing from the Fed through the discount window (supply of reserves). Changes in the rate charged on discount window loans would have affected the cost of obtaining reserves (price of reserves).

If changes in the borrowing by banks offset changes in reserves brought about by open market operations, then the quantity of reserves would be unchanged—the "scissors effect" noted above (figure 3 illustrates that there was an upward trend in the

[15]This channel is also quite similar to discussions about the market for reserves and implementation of monetary policy in the late 1990s/early 2000s (see Kroeger, McGowan, and Sarkar 2017).

[16]Changes in monetary gold should have had similar effects as open market operations, and we typically include such changes as a control variable (Friedman and Schwartz 1963 provide a detailed discussion of the amount of gold on the Fed's balance sheet.) In addition, many other changes in the Fed's balance sheet would affect reserves, with two other larger items being changes in Treasury balances held at the Fed or changes in "float" as check payments were credited to the receiving bank before the payments checks were debited from the paying bank. We also control for these in the regression analysis.

Figure 3. Reserves at the Federal Reserve Held by Commercial Banks

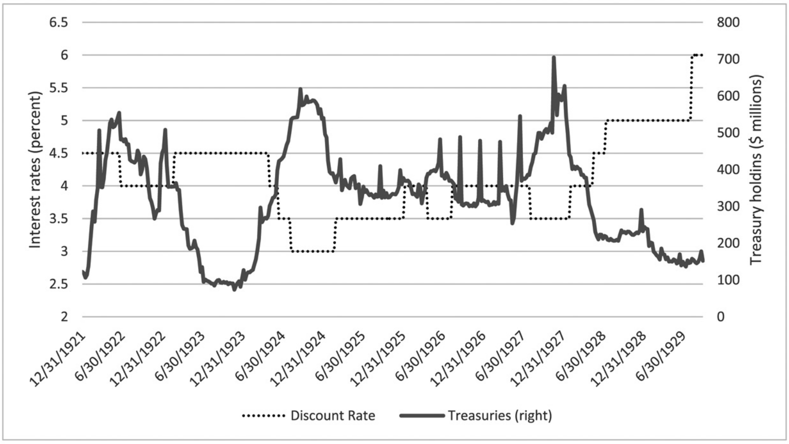


Source: Federal Reserve, Factors Affecting Bank Reserves and Condition Statement of Federal Reserve Banks (H.4.1 statistical release).

quantity of reserves in this period but no obvious cyclical fluctuations). To the extent that it was the quantity of reserves and its effects on monetary aggregates that mattered for the stance of monetary policy, that would imply that open market operations and changes in the discount and acceptance rate had limited effect on conditions.

Contemporary observers focused as much on the cost of obtaining reserves. Burgess (1936) argued that open market operations and interest rate policy could be used in a complimentary fashion, as seen in figure 4. For instance, when tightening policy, the Fed would conduct open market sales of Treasury securities to decrease the supply of reserves. This would make banks more reliant on obtaining reserves through the discount window. The Fed could then increase the discount window rate to raise the cost of using the window. Moreover, the Fed could also increase the discount on acceptance purchases so that banks received fewer reserves when they sold acceptances to the Fed which would further tighten the market for reserves. As the reserves market tightened, these pressures would have been passed on to other money markets. While there was not yet a liquid market in which banks could directly trade reserves—the federal funds market was still developing—banks could

Figure 4. Holdings of Treasury Securities and the Discount Rate



Source: Federal Reserve, Banking and Monetary Statistics, 1943.

trade reserves indirectly through trading securities and through their mutual participation in other overnight money markets and thereby transmit monetary policy to these other markets.

2.5.2 Portfolio Rebalancing

Another channel through which monetary policy might be transmitted to financial markets is the portfolio balance channel (see, for example, the discussion in Haldane et al. 2016). Underlying this mechanism is the idea that investors have particular preferences regarding the assets they hold. If the central bank were to purchase a particular type of asset, then the investors who had held that asset would purchase other assets with somewhat similar characteristics in ways that would change the relative prices of these assets. For instance, if certain investors have preferences for assets with a particular duration profile, then when the Fed purchases large amounts of longer-term Treasury securities, those investors might purchase large amounts of long-term corporate debt and compress the yield spreads between these asset classes. This compression in yield spreads might consequently promote issuance of corporate bonds.

In our period, the Fed purchased two types of assets—Treasury securities and bankers' acceptances. Under the portfolio balance channel, we might expect the effect on private money market rates from purchases of bankers' acceptances to be different, and possibly stronger, than the effects from purchases of Treasury securities, as the former are a more similar instrument in terms of credit risk and typical maturity.[17] Specifically, the acceptance rate in particular may have been able to influence financial market conditions in an additional, and more direct, way. Certain investors held both acceptances and other money market securities, such as commercial paper, as part of their portfolios. If the Fed changed the rate at which it bought acceptances, and thus changed the prevailing market rate, then we would expect the investors holding both assets to demand a different rate on other assets in their portfolios. Some contemporary observers mention such a dynamic (see Balabanis 1935). In this case, money market conditions should respond to changes in the acceptance rate to a greater degree than would be expected just from the changes in the Fed's holdings of acceptances and the associated change in the supply of reserves.[18]

However, this mechanism crucially depends on bankers' acceptances being treated as an active part of the portfolio. The Fed was heavily involved in this market and regularly held a considerable fraction of outstanding acceptances on its balance sheet. Moreover, the market price of acceptances seldom differed from the rate at which the Fed was willing to purchase them. Given this intervention, it is not certain that market participants would have used the

[17] The Fed purchased Treasury securities with a variety of maturities. We would expect that the portfolio balance effect of purchases of short-term certificates would have more impact on private rates than purchases of long-term bonds. However, newly issued certificates could still have maturities of up to one year. As the similarities between bankers acceptances and other private money market instruments are the clearest, we focus the analysis on acceptances.

[18] Our analysis of the portfolio rebalancing effect in the 1920s focuses on the effects of purchases of acceptances on other money market instruments where the effects are likely to be strongest. Studies of the more recent period have looked at the effects on longer-term assets, such as corporate bonds, which are more similar to the Treasury securities purchased by the Fed during the recent large-scale asset purchases (LSAPs). Thus, the mechanism being considered is the same, but the type of security that is the most natural comparison instrument is somewhat different.

rate on acceptances to price other assets, and the actions by the Fed may not have had any impact other than by changing the supply of reserves.

2.5.3 Signaling/Forward Guidance

One last mechanism by which policy might be transmitted is through signaling; by shaping expectations about whether monetary policy might be tighter or easier in the future, the Fed would be able to affect the prices of assets that would mature after the change in policy was expected to occur. Newspaper stories clearly indicate that market participants had expectations about actions the Fed might take. Additionally, newspaper stories and analysis of changes in the Fed's balance sheet and in the discount and acceptance rates make it clear that financial market participants were paying close attention to actions taken by the Fed. However the Fed did not provide much forward guidance.[19] Further, there is little discussion in the newspapers regarding how any actions taken by the Fed were reshaping the expectations of market participants about longer-term monetary policy.[20]

3. Impact of Monetary Policy Tools on Market Rates

In this section, we test whether changes in the monetary policy tools affected rates in markets, especially money markets. As part of this

[19] As policy tools were used in combination, it is possible that a sustained change in Treasury holdings might foreshadow a change in the discount rate or acceptance rate. To gauge this potential signaling effect of balance sheet movements, we looked to see whether there was indeed any relationship between a switch from accumulating/shedding Treasury holdings and the first decrease/increase in the discount rate by the New York Fed. We found that the first change in the discount rate occurred any time between 5 weeks and 24 weeks after the peak or trough in Treasury holdings. Given the wide range of time that could elapse between Treasury operations and interest rate operations, the role of any signaling effect on changes in market interest rates in any given week was likely pretty small.

[20] By contrast, Bordo and Sinha (2016) report notable discussion in newspapers of shifts in market expectations in response to changes in the Federal Reserve's balance sheet. They find that these shifts in expectations mattered for longer-term Treasury yields during 1932, when the Fed engaged in substantial purchases of Treasury securities.

first pass, we are agnostic about the channel through which the tools are operating. Before testing the transmission, we briefly review the important money markets of this period.

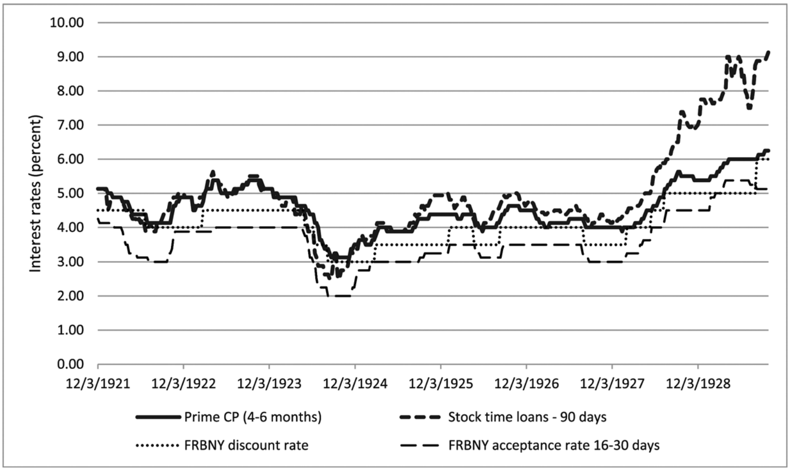
3.1 Review of Major Private Money Markets

There are two major private money markets that we use in our analysis. The prime commercial paper market was a money market in which moderately sized firms could borrow on an unsecured basis (larger firms would typically issue long-term bonds). The firms would issue short-term notes that would be purchased by a commercial paper house, which would arrange to distribute the paper around the country. While the houses would not guarantee the paper they sold, there were reputational consequences to selling paper that subsequently went bad.[21] Commercial paper was the most similar money market instrument to bankers' acceptances, as maturities tended to be similar and the underlying credit risk of both instruments was related to the condition of private nonbank firms. The most common buyers of commercial paper were banks, which would use this as a place to put funds on a short-term basis. This market was reasonably deep for much of the 1920s but started to fade by the end of that period.

The second money market is the market for brokers' loans, which were loans extended to New York City stockbrokers and brokerage houses and backed by equity securities traded on the New York Stock Exchange. This was the largest money market in the 1920s (see footnote 11). The New York City money center banks were prominent lenders, but other domestic banks, foreign banks, corporations, and investment trusts all provided funds. Broker loans were typically short-term demand loans, with a significant fraction of loans being overnight. Loans were typically rolled over, but could be called by the lender at any time, which resulted in the market being referred to

[21]Interest rates on this paper in the secondary market were determined by the quality of the firm issuing the paper. "Prime" commercial paper reflected the paper considered by the National Credit Office (a private firm affiliated with the R. G. Dun credit rating agency) to be of the highest quality, and the interest rates on that paper trading in the secondary market were used to construct the prime commercial paper rate.

Figure 5. Federal Reserve Policy Rates and Private Money Market Rates



Source: Federal Reserve, Banking and Monetary Statistics, 1943.
Note: FRBNY is the Federal Reserve Bank of New York.

as the "call loan market." However, a meaningful share (30 percent) of these loans were made on a time basis (90 days).

3.2 Setup of the Analysis

To get a general sense of the relationship between the policy tools and market rates, we plot two policy rates and two private money market rates at a weekly frequency in figure 5.[22] The two policy rates are the New York Fed's discount window rate and rate for purchases of acceptances with a maturity of 16–30 days. The private rates are the interest rate on four- to six-month commercial paper and the rate on time call loans (90 days). The figure suggests that, in general, policy rates moved in the same direction though they were not

[22] As noted by Wheelock (1991), the Fed was reportedly attentive to whether its policy rates were out of alignment with market rates. Thus, there is some possibility of reverse causality in that changes in market rates could have caused the Fed to adjust its policies. We address this concern below in section 3.4 using a dynamic framework that more explicitly allows for reverse causality and should further alleviate concerns about its impact on our findings.

adjusted at exactly the same time. From figure 4, we know that holdings of Treasury securities tended to move inversely with these rates. Figure 5 also indicates that the private market rates moved broadly at the same times as movements in the policy rates and the changes in holdings of Treasury securities. It is also clear from figure 5 that money market spreads during this time were positive and tended to move around over time, which is different than is the case now where these spreads tend to be fairly tight and steady over time.

To more formally test whether private rates moved in response to changes in policy instruments, we regress changes in private money market rates on the changes in the policy tools. Using changes is a better test than using levels because of reduced likelihood of spurious correlation and of issues associated with autocorrelation.[23] Specifically, we estimate the following equation:

$$r^{mm} = \alpha_0 + r^{ac'}\alpha_1 + r^{dw'}\alpha_2 + Treas'\alpha_3 + Gold'\alpha_4 + Bal'^{\alpha_5} + \alpha_m + \alpha_y + \varepsilon. \quad (1)$$

The vector r^{mm} denotes the changes in private market rates—the interest rate on four- to six-month commercial paper; time call loans (90 days); the interest rate on overnight, renewed call loans; and the interest rate on a long-term Treasury bond.[24] The matrices r^{ac}, r^{dw}, and $Treas$ capture the contemporaneous change and two lags of changes in the policy tools—the acceptance rate, the discount rate, and the Fed's holding of Treasury securities, respectively. The rate at which the Fed discounted acceptances varied by remaining maturity of the acceptance; we use the rate on acceptances with 16–30 days remaining maturity.[25] For Treasury operations, we scale the

[23] We also estimated the effects of the level of policy instruments on the level of market rates using an autoregressive distributed lag model; the long-run effects estimated using that approach are similar to those reported here using the first-difference specification.

[24] For most of the sample period these are bonds issued as part of the Fourth Liberty Loan program, which matured in 1938 (but were callable starting in 1933).

[25] The New York Fed had slightly different rates for acceptances in other maturity buckets. We use the 16- to 30-day rate, as it generally corresponds to the average maturity of acceptances bought by the New York Fed (see, for instance,

changes in the Fed's holdings to be in units of $100,000. Measuring changes in holdings of Treasury securities in dollar units facilitates estimating the size of the impact of Treasury purchases/sales; alternative scalings of the size of operations in Treasury securities produce qualitatively similar results. As indicated above, increases in the administered rates should raise private interest rates, while increased holdings of Treasury securities should reduce them.

Changes in the Fed's holdings of gold should also affect money rates in a manner similar to Treasury purchases. (Hanes 2006 finds that gold flows affected monetary conditions in the 1930s; gold inflows increased the supply of reserves and lowered longer-term Treasury yields.) To control for the impact of gold on financial conditions, we include contemporaneous and two-lag changes in the monetary gold (also measured in units of $100,000), denoted by the matrix *Gold*. We also control for other balance sheet developments—denoted *Bal*—that were outside of the control of the Fed, but which could affect reserves and short-term market conditions, such as changes in check float and in the Treasury's account at the Fed. (However, to keep the tables readable, we do not report the coefficients on these controls.) Seasonal factors may also be important, and we include monthly dummies in the regressions. We also include year effects.

All variables are measured at a weekly frequency. Our sample period is from June 1923 to August 1929, although for the overnight call loan market, we end the sample at the start of 1929, as this market appears to change in character during that year (see White 1990 for a detailed discussion).[26] We also omit a couple of months for the long-term bond near the end of the sample in which there was a change in the reference rate. Summary statistics of the variables used in the analysis are reported in table 2. Our regressions control for first-order serial correlation (although since we are looking at changes at a weekly frequency, the effects of serial correlation seem fairly minimal).

Federal Reserve 1923, table 62). Using other acceptance rates from the New York Fed produced similar results.

[26]This is also the time period when the Fed was using moral suasion to deter banks from using discount window loans to support "speculative excesses" in the

Table 2. Summary Statistics

Variable	Obs.	Mean	Standard Deviation	Min.	Max.
Weekly Variables					
Change in the Commercial Paper Rate (Basis Points)	314	.31	6.9	−37	25
Change in the Call Loan Time Rate (Basis Points)	314	1.2	15.2	−63	100
Change in the Call Loan Renewal Rate (Basis Points)	314	.68	74	−390	435
Change in the Long-Term Treasury Yield (Basis Points)	274	−.15	2.5	−9.0	12.0
Change in the New York Fed Discount Rate (Basis Points)	314	.47	10.9	−50	100
Change in the Discount at Which the New York Fed Purchased Acceptances with 16–30 Days to Maturity (Basis Points)	314	.36	7.7	−50	38
Change in Holdings of Treasury Securities ($100,000)	314	−.005	.25	−1.76	1.07
Change in Holdings of Acceptances ($100,000)	314	−.004	.20	−.65	.74
Change in Holdings of Gold ($100,000)	314	−.004	.19	−.67	.65
Monthly Variables					
Change in the Commercial Paper Rate (Basis Points)	83	.49	19.9	−74.7	44.4
Change in the New York Fed Discount Rate (Basis Points)	83	.60	23.6	−130	80
Change in the Discount at Which the New York Fed Purchased Acceptances with 16–30 Days to Maturity (Basis Points)	83	.14	19.7	−76.5	58.8
Change in Holdings of Treasury Securities ($100,000)	83	.0001	.05	−.12	.17
Change in Holdings of Gold ($100,000)	83	−.002	.04	−.12	.08
Change in Banks' Holdings of Acceptances ($1,000,000)	71	2.1	37.3	−121	115
Change in Bankers' Acceptances Outstanding ($100,000,000)	83	.11	.63	−1.3	2.7
Change in Industrial Production	81	1.1	.11	.90	1.42

Source: See data appendix.

3.3 Results

The results of the regression analysis are reported in table 3. They are, for the most part, in line with our expectations. We find that increases in both the acceptance rate and the discount window rate generally raised market interest rates (with the exception that overnight call loan rates are less responsive to the policy rates). The effects are strongest for commercial paper rates. The coefficients imply that an increase of 25 basis points in the acceptance rate would have increased the commercial paper rate by 9 basis points; a similarly sized increase in the discount window rate would have increased the commercial paper rate by 5 basis points. While these effects are economically meaningful, the pass-through from changes in policy rates to changes in market rates is notably less than the close to one-for-one pass-through observed in modern times. Increases in the discount window rate and the acceptance rate also appear to have lifted long-term rates. While the coefficient indicates that a 25 basis points increase in the acceptance rate or the discount window rate is associated with only a 2 basis points increase in the long-term Treasury rate in that week, the standard deviation of weekly changes in the long-term rate is only 2.5 basis points. Thus, while the effect is small, it is meaningful.[27]

Also as expected, increased holdings of Treasury securities are associated with reductions in private money market rates. We estimate that an increase of $500 million in the Fed's holding of Treasury securities (about the size of the increase in holdings of Treasury

stock market. Consequently, we would expect that changes in the discount rate would have less impact on call loan rates.

[27]We investigated alternative lag lengths. For most monetary policy variables, we are unable to reject that coefficients beyond a two-week lag are equal to zero. For changes in the acceptance rate, there is some evidence that lags up to four weeks may matter for the commercial paper rate, in which case the economic effects reported in table 3 would be somewhat larger, but lags beyond two weeks are not significant for other market rates. In the VAR specification below, we also investigated whether alternative lag lengths might matter. Different criteria suggest different things: Schwarz's Bayesian information criterion suggests that one lag is optimal; the Hannan and Quinn information criterion suggests that two lags is optimal, while Akaike's information criterion suggests four lags. Using different lag lengths, however, does not meaningfully change the economic magnitudes of the cumulative responsiveness of the variables to shocks. Thus, for consistency with the ordinary least squares (OLS) results, we continue to use two lags.

Table 3. Impact of Policy Instruments on Changes in Private Money Rates (weekly frequency)

Dependent Variable: Change in the Private Money Market Rates				
	Four- to Six-Month Prime CP	**Call Loans on Time (90-Day)**	**Call Loans Renewal (Overnight)**	**Long-Term Treasury**
Change in the Acceptance Rate				
Contemporaneous	.03	−.05	−.01	.04**
	(.05)	(.12)	(.33)	(.02)
Lagged One Week	.19***	.21*	−.06	.02
	(.05)	(.12)	(.35)	(.02)
Lagged Two Weeks	.14***	.07	.17	.01
	(.05)	(.12)	(.31)	(.02)
Change in the Discount Rate				
Contemporaneous	.08**	.09	.13	.04***
	(.03)	(.08)	(.25)	(.02)
Lagged One Week	.11***	.12	−.07	−.003
	(.03)	(.08)	(.24)	(.02)
Lagged Two Weeks	−.01	.02	.06	−.01
	(.03)	(.08)	(.24)	(.02)
Change in Treasury Holdings				
Contemporaneous	−.001	−.02	.03	−.01*
	(.02)	(.04)	(.11)	(.007)
Lagged One Week	−.03*	−.13***	−.26**	−.003
	(.02)	(.04)	(.11)	(.007)
Lagged Two Weeks	−.02	.01	−.10	−.003
	(.02)	(.04)	(.11)	(.008)
Change in Monetary Gold				
Contemporaneous	−.01	−.001	−.33**	.008
	(.02)	(.05)	(.14)	(.01)
Lagged One Week	−.05**	−.09*	−.52***	−.005
	(.02)	(.05)	(.14)	(.009)
Lagged Two Weeks	.01	−.11**	.33**	.005
	(.02)	(.05)	(.13)	(.009)
Include Float and Govt. Deposits	Yes	Yes	Yes	Yes
Month and Year Dummies	Yes	Yes	Yes	Yes
Intercept	−.02	−.06	−.27***	−.02
	(.02)	(.05)	(.09)	(.01)
AR1	−.12	.13	−.43	−.07
Observations	327	327	292	274
F-stat	3.9	1.8	7.0	1.7
Adjusted R^2	.24	.08	.41	.09

Source: Authors' calculations based on data described in the data appendix.
Notes: The symbols *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are in parentheses. Regressions adjust for first-order autocorrelation in the error terms (and we find no evidence of second-order serial correlation).

securities in 1924 when the Fed was easing policy and an amount equivalent to 3 percent of Treasury securities outstanding at the time) would have decreased the commercial paper rate by 250 basis points. Similarly, we estimate that such purchases would have resulted in an 80 basis point decline in the long-term Treasury rate. (However, there is considerable uncertainty around this estimate; the 95 percent confidence interval—i.e., two standard deviations—for the effect on the long-term Treasury rate ranges from 30 basis points to 130 basis points.)

We can compare this effect to the effect of purchases associated with the Fed's LSAPs in the wake of the financial crisis. Recent estimates (e.g., D'Amico and King 2013, Li and Wei 2013) suggest that the Fed's LSAPs starting on March 2009—which resulted in $300 billion in asset purchases (also roughly about 3 percent of outstanding Treasury securities)—resulted in a decline in Treasury yields of about 8–30 basis points.[28] Thus, our evidence suggests that Treasury purchases had larger effects in the 1920s than they do today. This result is similar to Bordo and Sinha (2016), who also find larger effects of Treasury purchases in the 1930s than in modern times. The larger impacts are consistent with the Treasury market being notably less liquid in the 1920s and the 1930s than today.

We also find that, again as expected, increases in monetary gold tended to reduce market interest rates. We find modest seasonal effects, with interest rates tending to rise a bit more, on average, during August, September, and October than other months. However the size of the seasonal effects are small (4 to 5 basis points), consistent with the Fed having successfully damped seasonal pressures (see Miron 1986; Carlson and Wheelock 2016).

3.4 Effects of Changes in Monetary Policy in a Dynamic System Which Takes the Gold Flows More Seriously

In the approach used above, we have assumed not only that the policy variables do not respond to week-to-week changes in

[28]The share of Treasury debt held by the Federal Reserve at the conclusion of the 1924 purchasing cycle was about 3 percent of total outstanding Treasury debt. At the conclusion of the Fed's large-scale purchases that was conducted following the March 2009 announcement, the Fed's holdings of Treasury debt were about 6.5 percent of the total outstanding Treasury debt.

market interest rates, but also that gold flows similarly do not respond to such fluctuations. However, it is possible, and indeed likely, that gold flows responded to changes in interest rates. In this case, an increase in the acceptance rate as the Fed tightened policy would raise the commercial paper rate or call market rate. The high rates would result in gold inflows.[29] Gold inflows would add to monetary gold on the Fed's balance sheet, generate an increase in reserves, and partly reverse the increase in policy rate. In addition, we also include a specification in which monetary policy responds to gold flows. Both of these two approaches should capture the interactions between monetary policy and gold flows that might matter for our analysis.

We explore these interactions using a structural vector autoregression (SVAR) consisting of five vectors in the following order: (i) changes in the holdings of Fed's Treasuries, (ii) changes in the acceptance rate, (iii) changes in the discount rate, (iv) changes in the Fed's holdings of gold, and (v) changes in the private commercial paper rate. Specifically, we estimate the following model:

$$\beta x_t = \Gamma_0 + \Gamma_1 x_{t-1} + \Gamma_2 x_{t-2} + \varepsilon_t, \tag{2}$$

where x_t is a matrix in which the columns are our five variables. To identify this model, we assume that changes in the policy interest rates and Treasury purchases are able to influence gold flows and market rates contemporaneously but market rates and gold flows do not have contemporaneous effects on the policy variables. We also allow gold flows and market rates to affect each other contemporaneously. In addition, all variables are allowed to affect each other with a one- and two-week lag. Accordingly, we impose the following structure on the matrix β:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ b_{g,t} & b_{g,r^{dw}} & b_{g,r^{ac}} & 1 & b_{g,r^{cp}} \\ b_{r^{cp},t} & b_{r^{cp},r^{dw}} & b_{r^{cp},r^{ac}} & b_{r^{cp},g} & 1 \end{bmatrix}.$$

[29] In the previous section, gold flows were treated as exogenous—for instance, due to changes in monetary policy abroad. This section allows gold flows to respond to changes in Fed policy.

Results of the SVAR involving the commercial paper rate as the market rate are shown in figure 6. We show the cumulative impacts of one-time, one-standard-deviation shocks to the different variables. (The size of the shocks thus matches the standard deviations of the variables as reported in table 2.) Results using other market rates instead of the commercial paper rate are similar.

We do observe some of the dynamics that motivated us to consider this framework. As shown in the next-to-the-last column, we find that gold inflows tended to reduce the commercial paper rate. Further, as shown in the last column, an increase in the commercial paper rate resulted in some gold inflows.

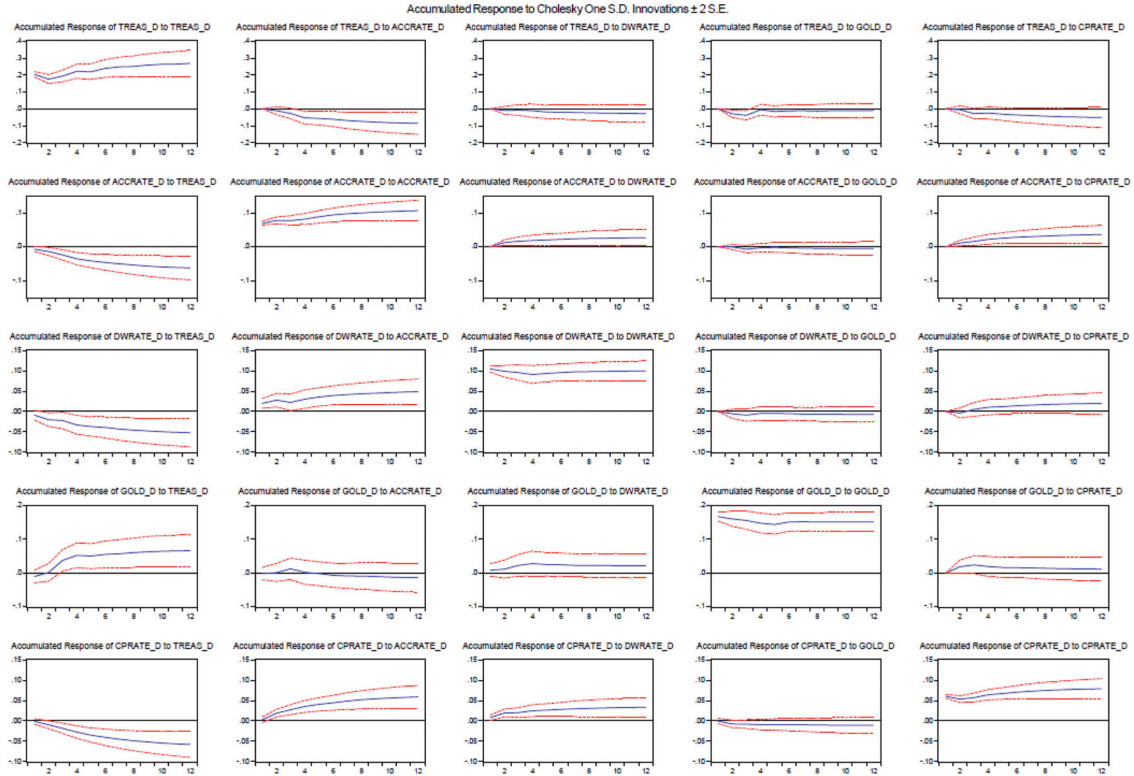
However, accounting for these dynamics does not diminish the impact of changes in policy rates on the commercial paper rate. As shown in the first three columns, the commercial paper rate continues to respond to the monetary policy variables of Treasury holdings, the acceptance rate, and the discount rate. If anything, the effects of a shock to the policy interest rates are even a little stronger than in the previous section. Here an 8 basis points increase in the acceptance rate is estimated to eventually result in a 6 basis point rise in the commercial paper rate and a 10 basis point increase in the discount rate increases the commercial paper rate by about 4 basis points.

The SVAR shown in figure 6 assumes that gold responds to policy contemporaneously but that policy does not respond to gold contemporaneously (policy may respond with a one-week lag). Some discussion of the 1920s argues that offsetting gold flows was an important part of the Fed's policy in the 1920s (Friedman and Schwartz 1963). In this case, a one-week lag may not sufficiently capture how policy was operating.

To more directly account for the possibility that monetary policy was sterilizing gold flows, we reorder the variables and adjust the restrictions in the SVAR to allow gold to affect the policy variables (and commercial paper rates) contemporaneously. The policy variables continue to affect the commercial paper rates contemporaneously. Commercial paper rates only affect other variables with a lag.

The results of this alternative SVAR are shown in figure 7. As before, we focus on cumulative responses. Here we observe that, while interest rate policy does not appear to have responded to gold

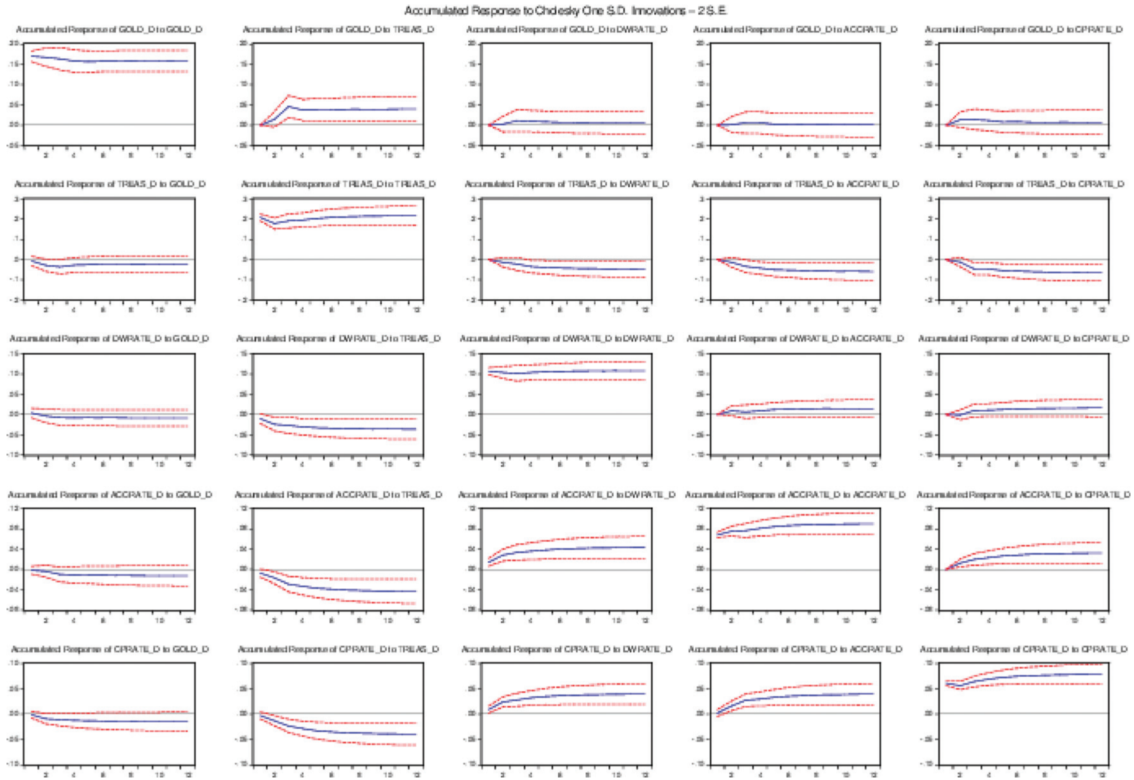
Figure 6. Vector Auto-Regression Results Using Commercial Paper Rate



Source: Federal Reserve, Factors Affecting Bank Reserves and Condition Statement of Federal Reserve Banks (H.4.1 statistical release), and Federal Reserve, Banking and Monetary Statistics, 1943.

Notes: The individual panels show the cumulative response over 12 weeks of each variable to shocks to each variable. The columns are arranged by the variable that provides the shock impetus. The first column shows the response of all variables to a shock to the Fed's holdings of Treasury securities. The remaining columns show, respectively, the response to shocks to the acceptance rate, the discount rate, the Fed's holdings of gold, and the commercial paper rate. The rows are arranged by which variable is responding to the shock. The top row shows the response in the Fed's holdings of Treasury securities to the various shocks. The remaining rows show, respectively, the response of the acceptance rate, the discount rate, the Fed's holdings of gold, and the commercial paper rate.

Figure 7. Vector Auto-Regression Results Using Commercial Paper Rate Allowing for Sterilization



Source: Federal Reserve, Factors Affecting Bank Reserves and Condition Statement of Federal Reserve Banks (H.4.1 statistical release), and Federal Reserve, Banking and Monetary Statistics, 1943.

Notes: The individual panels show the cumulative response over 12 weeks of each variable to shocks to each variable. The columns are arranged by the variable that provides the shock impetus. The first column shows the response of all variables to a shock to the Fed's holdings of Treasury securities. The remaining columns show, respectively, the response to shocks to the acceptance rate, the discount rate, the Fed's holdings of gold, and the commercial paper rate. The rows are arranged by which variable is responding to the shock. The top row shows the response in the Fed's holdings of Treasury securities to the various shocks. The remaining rows show, respectively, the response of the acceptance rate, the discount rate, the Fed's holdings of gold, and the commercial paper rate.

flows, an increase in gold does seem to have been associated with a decrease in Treasury holdings. This finding is consistent with at least partial sterilization.

In terms of our main object of interest, the response of the commercial paper rate to changes in the discount rate and the acceptance rate appear to be of about the same magnitude in this specification as in the one used to generate figure 6. We do find that the response of the commercial paper rate to a change in Treasury holdings is reduced slightly when we account for gold sterilization, but it is still economically and statistically meaningful.

4. Some Evidence Regarding Transmission Channels

The above analysis of the relationship between changes in the monetary policy tools and changes in the money rates was silent on the channels through which monetary policy mattered. In this section, we provide some evidence on which channels might be particularly relevant. We are especially interested in whether the monetary policy tools might have affected financial conditions beyond their effect on just the supply of reserves in the system. We begin by testing that hypothesis. Finding minimal effects of reserves, we then turn to a deeper analysis of the acceptance market and the unique role that the rate at which the Fed purchased acceptances and the acceptance market might have played to see if this sheds additional light on the channels through which monetary policy operated.

4.1 Testing the Reserves Channel

We start by considering the reserves channel and, in particular, whether changes in the discount rate and acceptance rated mattered beyond their impact on reserves. We do this by regressing changes in money rates directly on changes in reserves as well as changes in the acceptance rate and the discount rate. If the acceptance and discount rates continue to matter, then it ought to be because of a channel other than their impact on reserves. The particular regression specification we estimate is

$$r^{mm} = \alpha_0 + r^{ac'}\alpha_1 + r^{dw'}\alpha_2 + Reserves'\alpha_3 + \alpha_m + \alpha_y + \varepsilon, \quad (3)$$

which is quite similar to the previous specifications and continues to use contemporaneous and two lags of each of the independent variables.

The results are shown in table 4. We do not find that changes in reserves had much impact on money rates, other than that on the overnight call loan rate. This is consistent with some of the skepticism expressed about the reserve channel by Toma (1989).

Importantly, however, the effect of changes in the policy rates are similar to what we found above, even after controlling for reserves directly. The coefficients on changes in the acceptance rate and on the discount rate, for example, are quite close to those estimated in table 3. This suggests that monetary policy was effective in moving private money market rates but might have operated through a channel other than the reserves channel.

4.2 Other Channels of Transmission: What Can We Learn from the Acceptance Market?

We next focus on the acceptance market to see if this unique market might help shed some light on the channels through which monetary policy was being transmitted to private money markets. The Fed strongly supported the market for acceptances, as we discussed earlier. One way it did so was by encouraging banks to include acceptances in their portfolio of liquid assets in addition to commercial paper and call loans. And perhaps more importantly, the Fed acted as a major buyer in the market (as noted above, the Fed at times held more than 40 percent of outstanding acceptances). If the Fed was successful in convincing banks to hold acceptances in their portfolio of liquid assets (as was suggested by Balabanis 1935), then changes in the acceptance rate ought to have caused institutions holding acceptances to reprice other assets in their portfolio, and changing the acceptance rate should lead to larger effects than would be suggested just by the reserve channel. But, if the Fed distorted that market by being such a significant purchaser of these instruments that it became disconnected from other money markets, then we would expect that the reserve channel would account for nearly all the effect of this policy tool. To state this argument slightly more formally, if the reserves channel accounts for the entire transmission of the acceptance rate to market rates, then the estimated size of

Table 4. Impact of Policy Instruments on Private Money Rates (controlling directly for reserves)

Dependent Variable: Change in the Private Money Market Rates				
	Four- to Six-Month Prime CP	Call Loans on Time (90-Day)	Call Loans Renewal (Overnight)	Long-Term Treasury
Change in Reserves				
Contemporaneous	−.004	−.03	−.01	−.0004
	(.01)	(.02)	(.06)	(.004)
Lagged One Week	−.01	−.04	−.32***	.002
	(.01)	(.03)	(.07)	(.004)
Lagged Two Weeks	.01	.003	.02	.002
	(.01)	(.02)	(.06)	(.004)
Change in the Acceptance Rate				
Contemporaneous	.04	.05	.29	.05**
	(.05)	(.12)	(.37)	(.02)
Lagged One Week	.20***	.25**	−.32	.02
	(.05)	(.12)	(.39)	(.02)
Lagged Two Weeks	.14***	.11	.40	.01
	(.05)	(.12)	(.35)	(.02)
Change in Discount Rate				
Contemporaneous	.09**	.10	.22	.05***
	(.03)	(.08)	(.27)	(.02)
Lagged One Week	.11***	−.03	.08	−.003
	(.03)	(.08)	(.27)	(.02)
Lagged Two Weeks	−.01	−.10	−.15	−.02
	(.03)	(.08)	(.27)	(.02)
Month and Year Dummies	Yes	Yes	Yes	Yes
Intercept	−.03**	−.06	−.37***	−.02
	(.02)	(.04)	(.09)	(.01)
AR1	−.10	.11	−.28	−.08
Observations	327	327	304	274
F-stat	4.6	1.5	3.5	2.1
Adjusted R^2	.22	.04	.18	.09

Source: Authors’ calculations based on data described in the data appendix.
Notes: The symbols *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are in parentheses. Regressions adjust for first-order autocorrelation in the error terms (and we find no evidence of second-order serial correlation).

the effect on market rates from changing acceptance rate directly, as measured in equation (1), should be the same as estimating the effect of the change in the acceptance rate on changes in the Fed's holdings of acceptances (q^{ac}) and then looking at the effect of the changes in the Fed's holdings of acceptances on changes in money market rates, as measured in equations (3) and (4).[30]

$$q^{ac} = \alpha_0 + r^{ac'}\alpha_1 + r^{dw'}\alpha_2 + Treas'\alpha_3 + Gold'\alpha_4 + \alpha_m + \alpha_y + \varepsilon \tag{4}$$

$$r^{mm} = \alpha_0 + q^{ac'}\alpha_1 + r^{dw'}\alpha_2 + Treas'\alpha_3 + Gold'\alpha_4 + \alpha_m + \alpha_y + \varepsilon \tag{5}$$

(As above, the matrices r^{ac}, r^{dw}, and $Treas$ capture the contemporaneous change and two lags of changes in the policy tools—the acceptance rate, the discount rate, and the Fed's holding of Treasury securities, respectively.) If there are other channels at work besides the reserve channel, then the effect estimated in equation (1) should exceed the effect estimated in equations (4) and (5).

In addition, regardless of whether changes in the acceptance rate operated only through reserves or through additional channels, the impact of additional holdings of acceptances, Treasury securities, or gold should, dollar for dollar, have the same impact on money rates. This result is because changes in all of these instruments should have an equal effect on the supply of reserves.

We show the results of estimating equation (1) in table 3. Table 5 shows the results of estimating equation (4), and table 6 shows the results of estimating equation (5).[31] From table 3, we find that the impact of a 25 basis point increase in the acceptance rate would be expected to raise the commercial paper rate by 9 basis points. Based

[30] In fact, the effect of this two-step procedure should be slightly larger, as banks would presumably offset some of the change in reserves due to changes in sales of acceptances to the Fed with changes in the discount window borrowing so that the overall changes in reserves is less than the change in acceptances. That offsetting effect is captured in (1) but not in (3) and (4).

[31] We also used an SVAR to estimate this specification, similar to equation (2), but using acceptances on the balance sheet rather than the acceptance rate. The results regarding the impact of this balance sheet variable on the commercial paper rate are similar to the OLS specification in that an increase in holdings of acceptances has a small, negative effect on the commercial paper rate.

Table 5. Impact of Policy Instruments on Holdings of Acceptances (weekly frequency)

Dependent Variable: Change in Holdings of Acceptances	
	Change in Holdings of Acceptances
Change in the Acceptance Rate	
Lagged One Week	−.24*
	(.14)
Lagged Two Weeks	−.15
	(.14)
Lagged Three Weeks	−.08
	(.14)
Change in the Discount Rate	
Lagged One Week	.21*
	(.09)
Lagged Two Weeks	.10
	(.09)
Lagged Three Weeks	−.01
	(.09)
Change in Treasury Holdings	
Lagged One Week	−.11**
	(.05)
Lagged Two Weeks	−.02
	(.05)
Lagged Three Weeks	.08*
	(.05)
Change in Monetary Gold	
Lagged One Week	−.08
	(.06)
Lagged Two Weeks	−.13**
	(.06)
Lagged Three Weeks	.25***
	(.06)
Include Float and Govt. Deposits	Yes
Month and Year Dummies	Yes
Intercept	−.15***
	(.04)
AR1	−.13
Observations	327
F-stat	6.0
Adjusted R^2	.35

Source: Authors' calculations based on data described in the data appendix.
Notes: The symbols *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are in parentheses. Regressions adjust for first-order autocorrelation in the error terms (and we find no evidence of second-order serial correlation).

Table 6. Impact of Policy Instruments on Changes in Private Money Rates (weekly frequency)

Dependent Variable: Change in the Private Money Market Rates				
	Four- to Six-Month Prime CP	**Call Loans on Time (90-Day)**	**Call Loans Renewal (Overnight)**	**Long-Term Treasury**
Change in Holdings of Acceptances				
Contemporaneous	−.04**	.02	.63***	.00
	(.02)	(.05)	(.12)	(.01)
Lagged One Week	−.01	.01	−.11	−.00
	(.02)	(.05)	(.12)	(.01)
Lagged Two Weeks	−.01	−.09**	−.30**	−.02**
	(.02)	(.05)	(.12)	(.01)
Change in the Discount Rate				
Contemporaneous	.09***	.09	.09	.06***
	(.03)	(.08)	(.21)	(.02)
Lagged One Week	.15***	.15**	−.07	.01
	(.03)	(.08)	(.21)	(.02)
Lagged Two Weeks	.04	.06	.24	−.002
	(.03)	(.08)	(.21)	(.02)
Change in Treasury Holdings				
Contemporaneous	−.01	−.02	−.03	−.02*
	(.02)	(.04)	(.10)	(.01)
Lagged One Week	−.05**	−.14***	−.22**	−.01
	(.02)	(.04)	(.11)	(.01)
Lagged Two Weeks	−.03*	.02	−.07	−.004
	(.02)	(.04)	(.10)	(.01)
Change in Monetary Gold				
Contemporaneous	−.01	.004	−.36***	.01
	(.02)	(.50)	(.13)	(.01)
Lagged One Week	−.06**	−.09*	−.42***	−.01
	(.02)	(.05)	(.13)	(.01)
Lagged Two Weeks	−.01	−.11**	.26**	.001
	(.02)	(.05)	(.12)	(.01)
Include Float and Govt. Deposits	Yes	Yes	Yes	Yes
Month and Year Dummies	Yes	Yes	Yes	Yes
Intercept	−.03	−.07	−.18**	−.02
	(.02)	(.05)	(.09)	(.01)
AR1	−.06	.13	−.36	−.04
Observations	327	327	292	274
F-stat	2.8	1.8	8.6	1.7
Adjusted R^2	.16	.07	.47	.08

Source: Authors' calculations based on data described in the data appendix.
Notes: The symbols *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are in parentheses. Regressions adjust for first-order autocorrelation in the error terms (and we find no evidence of second-order serial correlation).

on table 5, we estimate that a 25 basis point increase in the acceptance rate would, over the course of three weeks, have reduced the amount of acceptances on the Fed's balance sheet by about $12,000. Based on the results reported in table 6, we estimate that these reductions in the Fed's holdings would have increased the commercial paper rate by 0.7 basis point. Thus, we find much stronger effects from the reduction in the acceptance rate than can be accounted for by the changes in reserves.

To test our additional hypothesis, we calculate the impact of a reduction in $12,000 of Treasury securities and of gold. Based on the estimates in table 3, we find that the amount of Treasury sales would have increased the commercial paper by 0.6 basis point. Similarly, that dollar-amount reduction to monetary gold would have raised the commercial paper rate by 0.6 basis point. Thus, we find evidence consistent with the idea that the same dollar-amount change in all three balance sheet items had about the same effect on the commercial paper rate.

4.3 Additional Evidence Involving Quantities of Acceptances

We are able to provide some additional evidence regarding the transmission of monetary policy and the interaction between the acceptance market and the commercial paper market using information on quantities. The portfolio balance channel of transmission suggests that when the Fed changes the rate at which it would purchase acceptances and hence the relative attractiveness of these instruments, that should affect investors' preferences for these instruments. We are able to test this using data on banks' holdings of acceptances. The portfolio balance channel also suggest that changes in the volume of these securities available to investors should affect the prices of similar securities. We test this by looking at whether changes in the issuance of acceptances affect the commercial paper rate.

In both cases the evidence is consistent with the portfolio balance channel. We only have a modest number of observations, as these data are only available monthly from 1925 onward, so the results should be treated with some caution. Nevertheless, it is useful to know that evidence that is available using this data is consistent with the portfolio balance channel.

4.3.1 Change in Holdings of Acceptances in Response to a Change in the Acceptance Rate

We first look at how banks' holdings of acceptances changed in response to a change in the rate at which the Fed would purchase acceptances. As shown in figure 1, when the Fed increased the acceptance rate, the market rate would advance by about the same amount and the rate of return that investors would earn by holding acceptances would increase. (This would apply to new purchases or originations. Investors would experience a mark-to-market loss on their current holdings of acceptances.) The opposite would occur when the Fed decreased rates. When the rate of return on acceptances increased, we would expect that investors and acceptance originators would seek to add more of these assets to their portfolio and shift away from holdings of securities that were close substitutes. The reduced demand for those similar securities would be expected to eventually reduce the price on those securities and raise the associated interest rates to be in line with those on acceptances.

In this section we test whether an increase (decrease) in the acceptance rate resulted in banks holding more (fewer) acceptances. As part of their monitoring of developments in the acceptance market, the Fed collected information on banks' holdings of acceptances on a monthly basis starting in 1925. We regress changes in this amount, measured in units of $1 million, on lagged changes in the acceptance rate as well as changes in other monetary policy tools (the discount rate and holdings of Treasury securities) and holdings of monetary gold. We also control for changes in economic activity by including changes in the index of industrial production and seasonal factors using month dummies. Given the small sample, controlling for serial correlation can bias the coefficients, so we report results using ordinary least squares and including an autoregressive term.

The results, reported in table 7, are consistent with the portfolio balance channel. An increase in the Fed's acceptance rate resulted in banks holding more of these securities, though after a one-month lag.[32] An increase in the acceptance rate of 25 basis points would

[32]There are various potential reasons for this lag. For instance, there may be issues associated with the mark-to-market accounting as the value of acceptances

Table 7. Impact of Acceptance Rate Change on Banks' Holdings of Acceptances

Dependent Variable: Change in Banks' Holdings of Acceptances		
	OLS	**AR1**
Change in Acceptances Rate		
Lagged One Month	85.1*	76.4*
	(43.1)	(39.4)
Lagged Two Months	−36.9	−21.8
	(34.6)	(32.4)
Change in the Discount Rate		
Lagged One Month	−6.5	−5.5
	(30.6)	(29.5)
Lagged Two Months	11.0	8.6
	(26.7)	(25.3)
Change in Holdings of Treasury Securities		
Lagged One Month	−198.7*	−191.2*
	(108.7)	(104.8)
Lagged Two Months	172.6	158.4
	(108.9)	(106.9)
Change in Holdings of Gold		
Lagged One Month	77.8	145.7
	(151.1)	(136.6)
Lagged Two Months	118.8	64.2
	(158.2)	(142.1)
Change in Industrial Production		
Contemporaneous	−24.6	−30.5
	(39.2)	(37.0)
Lagged One Month	−60.3	−62.7*
	(39.4)	(36.5)
Intercept	−14.2	−7.4
	(22.9)	(22.4)
Includes Month Dummies	Yes	Yes
AR1		−.31
Observations	55	55
F-statistic	1.4	1.5
R^2	.14	.16

Source: Authors' calculations based on data described in the data appendix.
Notes: The symbols *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are in parentheses. Regressions adjust for first-order autocorrelation in the error terms (and we find no evidence of second-order serial correlation).

have increased banks' holdings of acceptances by roughly $20 million, a change of about one-half standard deviation.

Banks' holdings of acceptances also appear to have been affected by other monetary policy tools as well as economic activity. For instance, when the Fed purchased more Treasury securities, which would have reduced interest rates, banks held fewer acceptances. We also find that banks tended to hold fewer acceptances on their balance sheets when industrial production was expanding more rapidly.

4.3.2 Impact of Changes in the Quantity of Acceptances on Commercial Paper Rates

The second supporting hypothesis we test is whether changes in the outstanding amounts of acceptances affected interest rates in the commercial paper market. If portfolio rebalancing and quantities of securities mattered, then we would also expect that changes in private market issuance of acceptances should also affect commercial paper rates. This effect is because increased issuance of acceptances creates additional supply of high-quality liquid assets. Since the Fed essentially pinned down the interest rates in the secondary market for acceptances, we should not expect that the price of these securities would change. However, we ought to see a change in the price of close substitutes, particularly commercial paper. (This reasoning assumes that the increased issuance represents a supply shock and that demand stays constant.)

We test this hypothesis by regressing the change in the commercial paper rate on changes in the amount of acceptances outstanding and the monetary policy variables. We expect that increases in acceptances should push up the commercial paper rate. We estimate this regression using monthly data owing to the frequency with which information is available regarding acceptances outstanding. We also control for changes in economic activity using changes in industrial production and seasonal factors using month dummies. Again we report results using ordinary least squares and including an autoregressive term.

on balance sheet securities declined. Alternatively, it may have taken some time for the change in secondary market rates, the ones affected by the Fed, to change primary market rates in such a way that banks would find it valuable to increase their holdings of newly originated acceptances.

The results, shown in table 8, indicate that increases in acceptances outstanding are associated with a rise in commercial paper rates, consistent with a portfolio balance channel. The effect occurs after two months, timing that is coincident with when banks increased their holdings of acceptances. (As noted above, we are assuming that changes in the supply of acceptances are a supply shock and that demand is constant. As we cannot be certain that demand is indeed constant, these results should be viewed as consistent with our hypothesis, but not as definitive proof of our hypothesis.) The monetary variables matter as before—increases in the acceptance rate and the discount rate pushed up commercial paper rates while increases in Fed holdings of monetary gold lowered them.

5. Conclusion

In this paper, we provide an overview of the three primary tools used by the Fed as part of the normal implementation of monetary policy in the 1920s—the discount window, purchases of bankers' acceptances, and purchases of government securities in the open market. These tools are fairly similar to the more "novel" or "unconventional" tools that were introduced by the Fed during the past decade. The large-scale asset purchases of Treasuries used by the Fed in the 1920s were of a smaller scale than the ones used by the Fed after the financial crisis, although operations during two easing cycles in the earlier period did more than triple the Fed's Treasury portfolio.

In addition, we show that the Fed was able to influence the private money market rates effectively with each of these tools. Indeed, the impact of changes in holdings of Treasury securities was larger in the 1920s than today. We also provide evidence that the rate at which the Fed purchased bankers' acceptances had quite substantial effects on other interest rates beyond what would be expected only due to the impact on acceptances sold to the Fed and the level of reserves. We find evidence consistent with a portfolio balance channel as banks appear to have changed their preferences for holding acceptances after the changes in the Fed's policy rate. Finding evidence that, even in the 1920s, the portfolio balance channel appears to have mattered provides a deeper understanding of the importance of different channels in the transmission of monetary policy. Taken

Table 8. Impact of Issuance of Acceptances on Commercial Paper Rates

Dependent Variable: Change in the Commercial Paper Rate		
	OLS	**AR1**
Change in Acceptances Outstanding		
Lagged One Month	−.02	−.03
	(.04)	(.04)
Lagged Two Months	.09**	.09**
	(.04)	(.04)
Change in the Acceptance Rate		
Contemporaneous	.51***	.52***
	(.09)	(.09)
Lagged One Month	.31***	.29**
	(.11)	(.11)
Change in the Discount Rate		
Contemporaneous	.16***	.16***
	(.05)	(.05)
Lagged One Month	−.08	−.07
	(.07)	(.07)
Change in Holdings of Treasury Securities		
Lagged One Month	−.08	−.05
	(.26)	(.27)
Lagged Two Months	.02	.02
	(.27)	(.26)
Change in Holdings of Gold		
Lagged One Month	−.70*	−.74*
	(.38)	(.38)
Lagged Two Months	.27	.34
	(.42)	(.41)
Change in Industrial Production		
Contemporaneous	.09	.10
	(.09)	(.09)
Lagged One Month	.12	.11
	(.10)	(.09)
Intercept	−.15***	−.15**
	(.05)	(.06)
Includes Month Dummies	Yes	Yes
AR1		−.10
Observations	54	54
F-statistic	15.8	17.7
R^2	.87	.88

Source: Authors' calculations based on data described in the data appendix.
Notes: The symbols *, **, and *** indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are in parentheses.

together, these results suggest that what is deemed as "unconventional" monetary policy during the recent decade can also be seen as fairly conventional policy used in ordinary times, and working through fairly conventional channels.

Appendix. Data

The weekly Fed balance sheet is from the H.4.1 statistical release available from the Federal Reserve Bank of St. Louis FRASER (Federal Reserve Archival System for Economic Research) website. The year-end 1926 balance sheet is from the Federal Reserve Board Annual Report for 1926.

Interest rates for commercial paper, new call loans, renewed call loans, time call loans, and the Federal Reserve Bank of New York policy rates are from the Federal Reserve's *Banking and Monetary Statistics 1914-1941.* Weekly data on long-term Treasury yields are from the *Federal Reserve Bulletin.*

Outstanding amounts of commercial paper and acceptances are from the Federal Reserve's *Banking and Monetary Statistics 1914-1941.*

The index of industrial production is from the Board of Governors of the Federal Reserve System.

References

Altavilla, C., G. Carboni, and R. Motto. 2015. "Asset Purchase Programmes and Financial Markets: Lessons from the Euro Area." ECB Working Paper No. 1864.

Balabanis, H. 1935. *American Discount Market.* Chicago: University of Chicago Press.

Beckhart, B. 1932. *The New York Money Market, Volume III: Uses of Funds.* New York: Columbia University Press.

Bordo, M., and A. Sinha. 2016. "A Lesson from the Great Depression that the Fed Might Have Learned: A Comparison of the 1932 Open Market Purchases with Quantitative Easing." NBER Working Paper No. 22581.

Bordo, M., and D. Wheelock. 2013. "The Promise and Performance of the Federal Reserve as Lender of Last Resort 1914-1933."

In *The Origins, History, and Future of the Federal Reserve: A Return to Jekyll Island*, eds. M. D. Bordo and W. Roberds, 59–98 (chapter 2). Cambridge University Press.

Burgess, W. R. 1936. *The Reserve Banks and the Money Market.* New York: Harper and Brothers Publishers.

Carlson, M., and B. Duygan-Bump. 2016. "The Tools and Transmission of Federal Reserve Monetary Policy in the 1920s." *FEDS Notes*, November 22. Board of Governors of the Federal Reserve System.

Carlson, M., and D. Wheelock. 2016. "Did the Founding of the Federal Reserve Affect the Vulnerability of the Interbank System to Systemic Risk?" Finance and Economics Discussion Series No. 2016-059, Board of Governors of the Federal Reserve System.

Carpenter, S., S. Demiralp, and Z. Senyuz. 2016. "Volatility in the Federal Funds Market and Money Market Spreads during the Financial Crisis." *Journal of Financial Stability* 25 (August): 225–33.

Clarida, R. 2014. "Navigating the New Neutral." *Economic Outlook*, PIMCO, November 13.

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, November): F415–F446.

D'Amico, S., and T. King. 2013. "Flow and Stock Effects of Large-Scale Treasury Purchases: Evidence on the Importance of Local Supply." *Journal of Financial Economics* 108 (2): 425–48.

Demiralp, S., B. Preslopsky, and W. Whitesell. 2006. "Overnight Interbank Loan Markets." *Journal of Economics and Business* 58 (1): 67–83.

Duffie, D., and A. Krishnamurthy. 2016. "Passthrough Efficiency in the Fed's New Monetary Policy Setting." Presentation at the Federal Reserve Bank of Kansas City Conference on "Designing Resilient Monetary Policy Frameworks," Jackson Hole, Wyoming, August 26.

Eichengreen, B., and M. Flandreau. 2012. "The Federal Reserve, the Bank of England, and the Rise of the Dollar as an International Currency, 1914-1939." *Open Economies Review* 23 (1): 57–87. doi: 10.1007/s11079-011-9217-1.

Federal Reserve. 1923. *Annual Report of the Federal Reserve Board.* Washington DC: Federal Reserve Board.

Ferderer, J. P. 2003. "Institutional Innovation and the Creation of Liquidity Financial Markets: The Case of Bankers' Acceptances, 1914-1934." *Journal of Economic History* 63 (3): 666–94.

Friedman, M., and A. Schwartz. 1963. *A Monetary History of the United States, 1867-1960.* Princeton, NJ: Princeton University Press.

Garbade, K. 2014. "Direct Purchases of U.S. Treasury Securities by Federal Reserve Banks." Staff Report No. 684, Federal Reserve Bank of New York.

Gertler, M., and P. Karadi. 2013. "QE 1 vs. 2 vs. 3. . . : A Framework for Analyzing Large-Scale Asset Purchases as a Monetary Policy Tool." *International Journal of Central Banking* 9 (Supplement 1, January): 5–53.

Goodfriend, M. 2014. "The Case for a Treasury-Federal Reserve Accord for Credit Policy." Testimony before the Subcommittee on Monetary Policy and Trade of the Committee on Financial Services, U.S. House of Representatives, Washington, DC, March 12.

Hackley, H. 1973. *A History of the Lending Function of the Federal Reserve Banks.* Washington DC: Board of Governors of the Federal Reserve System.

Haldane, A., M. Roberts-Sklar, T. Wieladek, and C. Young. 2016. "QE: The Story So Far." CEPR Discussion Paper No. 11691.

Hamilton, J. 1997. "Measuring the Liquidity Effect." *American Economic Review* 87 (1): 80–97.

Hanes, C. 2006. "The Liquidity Trap and U.S. Interest Rates in the 1930s." *Journal of Money, Credit and Banking* 38 (1): 163–94.

Holston, K., T. Laubach, and J. C. Williams. 2016. "Measuring the Natural Rate of Interest: International Trends and Determinants." Working Paper No. 2016-11, Federal Reserve Bank of San Francisco.

Kiley, M. T. 2018. "Quantitative Easing and the 'New Normal' in Monetary Policy." Finance and Economics Discussion Series No. 2018-004, Board of Governors of the Federal Reserve System. https://doi.org/10.17016/FEDS.2018.004.

Krishnamurthy, A., A. Vissing-Jorgensen, S. Gilchrist, and T. Philippon. 2011. "The Effects of Quantitative Easing on Interest

Rates: Channels and Implications for Policy." *Brookings Papers on Economic Activity* (Fall): 217–87.

Kroeger, A., J. McGowan, and A. Sarkar. 2017. "The Pre-Crisis Monetary Policy Implementation Framework." Staff Report No. 809, Federal Reserve Bank of New York.

Li, C., and M. Wei. 2013. "Term Structure Modeling with Supply Factors and the Federal Reserve's Large-Scale Asset Purchase Programs." *International Journal of Central Banking* 9 (1, March): 3–39.

Meltzer, A. 2003. *A History of the Federal Reserve, Volume 1, 1913-1951.* Chicago: University of Chicago Press.

Miron, J. 1986. "Financial Panics, the Seasonality of the Nominal Interest Rate, and the Founding of the Fed." *American Economic Review* 76 (1): 125–40.

New York Clearing House Association. 1920. *Constitution of the New York Clearing House Association. Also Clearinghouse Rules, Scale of Fines, Collection Charges, and Holiday Laws of New York State.* New York: New York Clearing House Association.

Reifler, W. 1930. *Money Rates and Money Markets in the United States.* New York: Harper & Brothers Publishers.

Reis, R. 2016. "Funding Quantitative Easing to Target Inflation." Presentation at the Federal Reserve Bank of Kansas City Conference on "Designing Resilient Monetary Policy Frameworks," Jackson Hole, Wyoming, August 26.

Small, D. H., and J. Clouse. 2005. "The Scope of Monetary Policy Actions Authorized under the Federal Reserve Act." *Topics in Macroeconomics* 5: 1.

Stein, J. 2012. "Monetary Policy as Financial Stability Regulation." *Quarterly Journal of Economics* 127 (1): 57–95.

Toma, M. 1989. "The Policy Effectiveness of Open Market Operations in the 1920s." *Explorations in Economic History* 26 (1): 99–116.

Turner, B. 1931. *The Federal Funds Market.* New York: Prentice-Hall, Inc.

Wheelock, D. 1991. *The Strategy and Consistency of Federal Reserve Monetary Policy, 1924-1933.* Cambridge: Cambridge University Press.

White, E. 1990. "The Stock Market Boom and Crash of 1929 Revisited." *Journal of Economic Perspectives* 4 (2): 67–83.

Wicker, E. 1966. *Federal Reserve Monetary Policy 1917-1933.* New York: Random House.

Yellen, J. 2016. "The Federal Reserve's Monetary Policy Toolkit: Past Present, and Future." Remarks at the Federal Reserve Bank of Kansas City Conference on "Designing Resilient Monetary Policy Frameworks," Jackson Hole, Wyoming, August 26.

Policy and Macro Signals from Central Bank Announcements*

Paul Hubert[a] and Becky Maule[b]
[a]Banque de France and Sciences Po – OFCE
[b]Bank of England

How do private agents understand central bank actions and communication? This paper investigates private agents' interpretation of central bank signals about future policy and the macroeconomic outlook that are conveyed by both policy decisions and macroeconomic communication. Using U.K. data, we assess the effects of these two types of central bank announcements on inflation swaps and stock returns at daily and monthly frequencies. We find that policy decisions convey signals about the macroeconomic outlook at the daily frequency, but that policy signals dominate at the monthly frequency such that asset prices respond negatively to contractionary policy, consistent with the usual transmission mechanism. We also find that inflation expectations respond positively to the Bank of England's macroeconomic information surprises, consistent with agents taking a signal about the economic outlook.

JEL Codes: E52, E58.

*We thank Christophe Blot, James Cloyne, Camille Cornand, Jérôme Creel, Martin Ellison, Rodrigo Guimarães, Refet Gürkaynak, Stephen Hansen, Frédéric Jouneau-Sion, Michael McMahon, Silvia Miranda-Agrippino, Jon Relleen, Giovanni Ricco, Garry Young, and seminar participants at the GATE, Bank of England, OFCE, European Central Bank, Deutsche Bundesbank, CEPII, Banque de France, the Workshop on Empirical Monetary Economics 2015, the Workshop on Probabilistic Forecasting and Monetary Policy, GDRE 2015, the Central Bank Design workshop of the 2015 BGSE Summer Forum, the 2015 World Congress of the Econometric Society, and AFSE 2015 for helpful discussions and comments at different stages of this project. We also thank Simon Strong for research assistance. Paul Hubert thanks the Bank of England for its hospitality. Any remaining errors are our own responsibility. Any views expressed are solely those of the authors and so cannot be taken to represent those of Banque de France, Bank of England, or members of the Monetary Policy Committee or Financial Policy Committee. Corresponding author: paul.hubert@sciencespo.fr. Address: 31 rue Croix des Petits Champs, 75001 Paris, France.

1. Introduction

Private agents' interpretation of monetary policy decisions and central bank communication is central to the formation of their beliefs and therefore to the transmission of monetary policy. The management of private expectations has thus become a key feature of central banking (Woodford 2005). In this respect, this paper proposes to investigate the effect of two different types of central bank announcements, one about the policy decision and the other about policymakers' assessment of the macroeconomic outlook.

In a framework with perfect information, private agents' and policymakers' expectations are aligned. So policy decisions only reveal monetary innovations, i.e., deviations from the systematic response of policy to the macroeconomic outlook. However, in a setup with different information sets (i.e., the central bank and the private sector do not share the same information set), the policy decision can convey information about changes in both policymakers' preferences and their view of future macroeconomic developments. For example, an increase in the policy rate could signal to private agents that an inflationary shock will hit the economy in the future, causing higher inflation. Alternatively, the same increase in the policy rate may be interpreted as a simple contractionary monetary shock, which will lead to lower inflation in the future. If the first interpretation is given more weight, then increasing the policy rate will lead to higher private inflation expectations, whereas if the second is given more weight, then increasing the policy rate will decrease private inflation expectations. The former effect has been labeled the "signaling channel" of monetary policy. Romer and Romer (2000) find evidence of the Federal Reserve's decisions revealing its private information about the state of the economy.[1] Ellingsen and Söderström (2001) establish that the private response to policy decisions may reflect a mix of the responses to the monetary innovation and to

[1] D'Agostino and Whelan (2008), Gamber and Smith (2009), Hubert (2015a, 2015b), and Rossi and Sekhposyan (2016) have challenged the result that central banks have a forecasting advantage relative to the private sector. Recent evidence suggests that the advantage has vanished in the most recent period. This result, however, does not rule out the fact that policy decisions reveal the central bank information set. Both information sets could be different and yield to similar forecasting performance.

the macroeconomic information conveyed by the policy instrument. Campbell et al. (2012, 2017), Tang (2015), Melosi (2017), and Nakamura and Steinsson (2018) complement the analysis of this information channel.[2]

The same issue applies to central bank macroeconomic communication. For instance, an increase in a central bank's inflation projections could signal a future inflationary shock, causing higher inflation expectations. Alternatively, because policy decisions are informed by the economic outlook, and the central bank might respond to developments which raise inflation by tightening policy to bring it back toward its target, the same increase in central bank inflation projections could be interpreted as a signal about a future policy tightening, leading to lower expected inflation. Hence, the central bank's projections can also send signals about its view of macroeconomic developments to private agents, influencing their beliefs about the economic outlook, as well as about the outlook for policy. We define the former channel as a "macro outlook signal" and the latter channel as a "policy signal."[3]

If a positive signal about the macro outlook is taken from either a policy decision or a change in central bank economic projections, inflation expectations will increase, whereas if either tighter policy or higher economic projections are taken to signal a contractionary policy shock, inflation expectations will decrease. Which channel—the macro outlook or policy signal—dominates matters, given that the effects of monetary policy depend on how private agents interpret changes in the policy rate or in central bank projections.

The contribution of this paper is to examine empirically which of those interpretations is given more weight and to document the

[2]This effect may be one of the reasons for the positive response of inflation to monetary shocks documented in the vector autoregression (VAR) literature as the "price puzzle" (Sims 1992). Castelnuovo and Surico (2010) finds that including inflation expectations in VARs removes this price puzzle. The signaling issue has also received attention from a theoretical perspective. See Angeletos, Hellwig, and Pavan (2006), Walsh (2007), Baeriswyl and Cornand (2010), and Kohlhas (2014).

[3]We use the term "policy signal" for the classical monetary transmission channel and the term "macro outlook signal" for what Melosi (2017) calls the "signaling channel" of monetary policy. That is because we study the information content of a central bank's macroeconomic projections, so the usual terminology is not appropriate.

importance of the signals that both policy decisions and central bank macroeconomic projection surprises send about the macroeconomic or the policy outlook. To do so, we assess, for the United Kingdom, whether and how market-based inflation expectations and stock prices respond to policy decisions and to the publication of the Bank of England's Inflation Report, which contains its macroeconomic projections.[4] The sign of the estimated effects of policy decisions and macroeconomic projections is indicative of the relative weight private agents put on each signal.

To estimate private agents' responses to central bank announcements, we make use of two features of the U.K. data. First, the fact that policy decisions and the Inflation Report (IR) containing the Bank of England (BoE)'s macroeconomic projections were released on different days until August 2015 enables us to carefully measure the surprise component of the two types of releases using daily data.[5] Second, a necessary requirement for identification is that the information set revealed by the central bank (such as its projections) is not a function of the current policy decision, so both monetary surprises and central bank information surprises can be separately identified.[6] We exploit the fact that the BoE publishes macroeconomic projections that are conditioned on the path for the policy instrument implied by financial market interest rates prior to the policy meeting.[7] As these projections are not conditioned on the

[4]It is worth stressing that the focus of this paper is on the effects of the release of central bank macroeconomic information, not communication about the future likely path of the policy rate, the forward guidance policy (see, e.g., Campbell et al. 2017; Andrade et al. 2019), or whether communication is relatively more hawkish or dovish (see, e.g., Ehrmann and Fratzscher 2007; Rosa and Verga 2007).

[5]After August 2015, both are released simultaneously at 12:00, so even intraday data would not enable us to do so.

[6]This paper focuses on quantitative communication, the release of central bank projections, and abstracts from the issue of quantifying qualitative communication like statements or minutes (see Hansen and McMahon 2016 for this kind of exercise, and Hubert 2017 for a comparison of the effects of both types of communication).

[7]For comparison, Federal Open Market Committee (FOMC) projections are conditioned on FOMC members' views of "appropriate monetary policy" which corresponds to the future interest rate path that best satisfies the Federal Reserve's mandate.

BoE's policy decision, it enables us to identify separately projection surprises and monetary innovations.

Our empirical analysis proceeds in two steps. First, we use a daily-frequency event-study analysis to provide a causal inference of the effect of both monetary policy decisions and the publication of central bank information on inflation swaps and stock prices. Following Hanson and Stein (2015), we use the daily change in two-year nominal interest rates around announcements to measure monetary and IR surprises. In the spirit of Cieslak and Schrimpf (2019) and Jarocinski and Karadi (2020), we exploit the fact that policy and macro signals have similar predictions for interest rates but different predictions for equity prices, such that the co-movement of these variables on the day of policy announcements should be informative of which signal, about policy or the state of the economy, dominates. Second, we estimate the effects of exogenous shocks to policy and to Bank's inflation and output projections at the monthly frequency in a framework derived from the information frictions literature, controlling for the contribution of inflation surprises, news shocks, and changes in private output and interest rate expectations.[8] We follow the identification strategy of Romer and Romer (2004) applied to U.K. data by Cloyne and Huertgen (2016). Following Blanchard, L'Huillier, and Lorenzoni (2013) and Miranda-Agrippino and Ricco (2020), we augment the original approach with the information set of private agents to take into account the influence of potential different information sets on the identification problem.

We find that private inflation expectations and stock prices respond positively to contractionary monetary surprises on average at the daily frequency, such that the dominant signal of policy decisions is about the macro outlook. However, in a monthly frequency framework, inflation expectations and stock prices respond negatively to contractionary monetary shocks. The signaling channel might still be at work in dampening the negative response of inflation expectations, but it is dominated by the effect of policy

[8]The use of inflation swaps in levels calls for correcting for potential term, liquidity, and inflation risk premiums. To do so, we use a regression-based approach following Gürkaynak, Levin, and Swanson (2010), Gürkaynak, Sack, and Wright (2010), and Soderlind (2011) described in section A.1 of the online appendix (available at http://www.ijcb.org).

signals. One potential explanation for the differentiated results is that differences in information sets are more pronounced at higher frequencies so private agents rely more on macro signals from policy actions. However, at lower frequencies, private agents are able to acquire more information, consistent with sticky-information models (see Mankiw and Reis 2002), such that the signaling effects are smaller.

With respect to the release of central bank information, we find that inflation expectations respond positively to positive IR surprises at the daily frequency, such that the dominant signal is about the macro outlook. The result holds at the monthly frequency with central bank inflation projections. Overall, that suggests that private agents put more weight on the signal that the central bank information set conveys about future economic developments than about the policy outlook, in contrast to the predictions of full-information models.

One potential concern may relate to the fact that the policy has reached the effective lower bound (ELB) over the sample studied here, and that monetary policy has taken various dimensions in the meantime. We attempt to circumvent this issue in various ways. First, the daily-frequency setup measures private responses to all policy decisions (conventional and unconventional) on the announcement day. In addition, using two-year spot nominal interest rates, following Hanson and Stein (2015), enables us to capture news about the future policy path. Second, in the monthly-frequency setup, we use a shadow rate that encompasses the overall stance of monetary policy. Third, we estimate the robustness of our results on subsamples before and after the ELB. Another potential concern is that the estimated effect of signals could be unrelated to the central bank information set and could instead reflect other macroeconomic news. If so, our estimates would suffer from an omitted-variable bias. To address this concern, we modify our baseline daily frequency model and control for the news surprises in six of the most important macroeconomic data releases, such as inflation, PMI, industrial production, and earnings. At the monthly frequency, we attempt to control for as much information as possible: we include various measures of slack, inflation, and sentiment in addition to inflation surprises, news indexes, and other private macro forecasts.

This work is related to different strands of research about the role of central bank communication in policymaking (see, e.g., Woodford 2005; Blinder et al. 2008; Reis 2013), its effects on inflation expectations (see, e.g., Gürkaynak, Sack, and Swanson 2005a; King, Lu, and Pasten 2008; Pedersen 2015; Lyziak and Paloviita 2018) or their dispersion (see, e.g., Fujiwara 2005; Ehrmann, Eijffinger, and Fratzscher 2012; Hubert 2014), or how it may help in predicting future policy decisions (see, e.g., Jansen and De Haan 2009; Hayo and Neuenkirch 2010; Sturm and De Haan 2011).[9]

The rest of the paper is organized as follows. Section 2 describes our framework, section 3 the event-study estimation, section 4 the monthly-frequency analysis, and section 5 concludes.

2. Framework

2.1 Some Theoretical Predictions

First, we derive predictions for the expected effects of monetary policy and central bank projections on private inflation expectations based on a standard macroeconomic framework, such as a three-equation New Keynesian model. In such a framework where private agents have full information, so central bank projections and private expectations are aligned, contractionary monetary policy has a negative effect on private inflation expectations, through the usual transmission channels. Increases in central bank projections at a given horizon have a negative effect on private inflation

[9]This paper also refers to a large literature focusing on the expectation formation process departing from the full-information rational expectation hypothesis to account for some empirical regularities about the persistence of private expectations, with among others Evans and Honkapohja (2001), Bullard and Mitra (2002), Mankiw and Reis (2002), Sims (2003), Branch (2004, 2007), Orphanides and Williams (2005), or Mackowiak and Wiederholt (2009). Another strand of the literature tries to explain macroeconomic outcomes with expectations (see, e.g., Nunes 2010 and Adam and Padula 2011), while various empirical works focus on the characteristics, responsiveness to news, dispersion or anchoring of expectations (see, e.g., Swanson 2006; Capistran and Timmermann 2009; Crowe 2010; Gürkaynak, Levin, and Swanson 2010; Beechey, Johannsen, and Levin 2011; Coibion and Gorodnichenko 2012, 2015; Hubert 2014, 2015a; Ehrmann 2015).

expectations further ahead, because central bank projections enter the central bank reaction function and are interpreted as signals about future policy reactions: higher expected future inflation leads agents to anticipate higher future nominal interest rates, especially when the policy rate exhibits persistence. In this framework, neither policy decisions nor central bank projections convey signals about the macroeconomic outlook.

Second, we derive predictions for the expected effects of monetary policy and projection surprises under a framework with information frictions. That assumption is consistent with works by Coibion and Gorodnichenko (2012, 2015) and Andrade and Le Bihan (2013), which provide empirical evidence of rejection of full-information models.[10] In a framework with different information sets, we assume the central bank sets its interest rate i_t as a function of its own inflation, $\pi_{t,h}^{CB}$, and output, $x_{t,h}^{CB}$, projections for horizon h, and potentially some other macro variables, ω_t:

$$i_t = f(i_{t-1}, \pi_{t,h}^{CB},\ x_{t,h}^{CB}, \omega_t) + \varepsilon_t^i, \tag{1}$$

and where ε_t^i is the monetary innovation, capturing policymakers' deviations from their policy rule, and which is orthogonal to central bank inflation and output projections and to other macro variables in ω_t.

In this setup where private agents and the central bank have different information sets, when the observed policy rate differs from private agents' policy expectations, private agents would not be able to infer whether the central bank has changed its *own* view of future inflation and output or whether there has been a monetary shock. Changes in the policy rate may therefore convey signals about both future macroeconomic developments and the policy stance to private agents. Private agents face a multidimensional signal-processing problem: they could take either of two signals—one about macro developments and one about future policy—from one observable

[10] Rational expectation models with information frictions such as in Woodford (2001), Mankiw and Reis (2002), and Sims (2003) highlight how departing from the assumption of full information can account for empirical patterns about inflation expectations as well as leading to different policy recommendations.

variable. Said differently, private agents can misperceive changes in policy or projections for a mix of shocks in the economy, which gives room for macro or policy signals—as modeled by Melosi (2017). The same reasoning applies to surprises to central bank projections. Depending on the information set of private agents, central bank projections could either convey a signal to private agents about a change in the central bank's view of future inflation and output or convey a signal about future policy through the reaction function.

Policy and macro signals are expected to have different implications for private inflation expectations. If either a higher policy setting or higher projections are taken to signal a contractionary policy shock (i.e., the policy signal dominates the macro outlook signal), inflation expectations will decrease. In contrast, if a positive signal about the macro outlook is taken from either a policy decision or a change in the central bank's projections (i.e., the macro outlook signal outweighs the policy signal), inflation expectations will increase.

2.2 The Empirical Strategy

The literature commonly uses a high-frequency event-study analysis to estimate the effect of monetary policy on asset prices. The key assumption is that the reaction of asset prices that are continually affected by various factors can be specifically attributed to monetary news on the day of the policy announcement. If one assumes that the central bank (CB) and private agents (PA) have different forecasts (E[.]) about the state of the economy (Y_t), but private agents know the specification of the reaction function, then the signaling problem can be written as follows:

$$\begin{aligned} i_t &= \varphi \mathrm{E}^{CB}[\mathrm{Y}_t] + \varepsilon_t^i \\ \mathrm{E}^{\mathrm{PA}}[i_t] &= \varphi \mathrm{E}^{\mathrm{PA}}[\mathrm{Y}_t] \\ mps_t &= i_t - \mathrm{E}^{\mathrm{PA}}[i_t] = \varepsilon_t^i + \varphi(\mathrm{E}^{\mathrm{CB}}[\mathrm{Y}_t] - \mathrm{E}^{\mathrm{PA}}[\mathrm{Y}_t]), \end{aligned} \tag{2}$$

where i_t is the realized policy rate, $\mathrm{E}^{\mathrm{PA}}[i_t]$ is the expected policy rate, and mps_t is the monetary policy surprise. In this setup, the monetary policy surprise mps_t is a linear combination of the monetary policy shock ε_t^i and the information shock ε_t^Y which corresponds to the difference between the information sets of the central bank and private agents. Said differently, the monetary policy surprise can be seen as a shock to the information set of private agents. The most recent literature (Cieslak and Schrimpf 2019, Jarocinski and Karadi 2019) has exploited the fact that ε_t^i and ε_t^Y have the same theoretical prediction for interest rates but different predictions for equity prices, such that the co-movement of these variables on the day of the policy announcement should be informative about which signal—about policy or the state of the economy—dominates. One compelling feature of the Bank of England's setup is that Monetary Policy Committee (MPC) decisions and the IR were not published on the same day until August 2015, so it is possible to assess the interpretation given by private agents to each announcement separately.[11]

Another approach in the literature to identify the causal impact of monetary policy measures changes in the policy instrument (i_t) that are unrelated to macroeconomic conditions and the information set of policymakers ($\mathrm{E}^{\mathrm{CB}}[\mathrm{Y}_t]$). The residuals from that estimation (ε_t^i) are taken to represent monetary policy shocks, i.e., exogenous shocks to the policy instrument. We make use of a specific feature of the Bank data to test our research question. We exploit the fact that the Bank of England publishes macroeconomic projections that are conditioned on the path for the policy instrument implied by financial market interest rates prior to the policy meeting, rather than a preferred interest rate path of the MPC, so do not contain the effect of the policy decision.[12] As these projections are not conditioned on the Bank's policy decision, we are able to separately identify projection surprises and monetary shocks.

[11]The IR was published on average four business days after the MPC meeting, with a minimum of two days on May 2005, May 2010, and May 2015 and a maximum of five days in February 2015. From August 2015 and following Warsh (2014)'s report, the IR started to be published at the same time as policy decisions.

[12]For comparison, FOMC projections are conditioned on FOMC members' views of "appropriate monetary policy" which corresponds to the future interest rate path that best satisfies the Federal Reserve's dual objectives of maximum employment and price stability.

3. The Effect of Monetary and IR Surprises at the Daily Frequency

3.1 The Overall Effect of Monetary and IR Surprises

We use standard monetary surprises that have been extensively used in the literature to assess the effect of unexpected policy decisions on asset prices. We apply the same methodology to measure surprises related to the publication of the IR. We thus augment the regression with an additional term to assess the effects of the publication of the central bank macroeconomic information set. The literature relies on the following regression:

$$\Delta \mathrm{y}_t = \alpha + \beta_1 \mathrm{MPC}_t + \beta_2 \mathrm{IR}_t + \varepsilon_t, \tag{3}$$

where MPC_t denotes the surprise component of the policy decision announced by the MPC, IR_t captures the surprise component of the macroeconomic information published in the IR, $\Delta \mathrm{y}_t$ denotes the change in the asset price considered over an interval that brackets the monetary policy announcement, and ε_t is a stochastic error term that captures the effects of other factors that influence the asset price in question.

The literature commonly uses a high-frequency event-study analysis to estimate equation (3), which cannot be estimated with monthly or quarterly data due to reverse causality and omitted-variables bias. The measured effect of monetary policy on asset prices could easily capture the response of monetary policy to earlier changes in asset prices in the month or quarter. In addition, changes in monetary policy and asset prices could respond to macroeconomic news released during the month or quarter. Using higher-frequency data and a tight window around the policy decision enables us to address these two issues. The key assumption is that the reaction of asset prices that are continuously affected by various factors can be specifically attributed to monetary news on the day of the policy announcement. Since asset prices adjust in real time to macroeconomic news, their movements during the window of a policy announcement only reflect the effect of news about monetary policy. This is crucial for identification since it strips out the endogenous variation in asset prices associated with other shocks than monetary news. Using daily data, Cook and Hahn (1989), Kuttner (2001), Cochrane and Piazzesi (2002), or Faust, Swanson, and Wright (2004) have initiated this approach.

A large consensus has formed about the content of monetary policy news: the main piece of information on central bank announcement days relates to changes in the future likely policy path. Following Gürkaynak, Sack, and Swanson (2005a), Campbell et al. (2012), and Hanson and Stein (2015), our identification strategy is based on the idea that a primary share of the news contained in MPC announcements is about the expected path of future policy (whether it is the policy rate during a period of conventional monetary policy or asset purchases in the most recent period) over the next several quarters as opposed to surprise changes in the current policy stance. A simple and transparent way to capture revisions to the expected path of policy over a given horizon is to use the daily change in the nominal gilt yield at this horizon on MPC announcement dates as our proxy for monetary policy news. Following Hanson and Stein (2015), we use interest rates at the two-year maturity to measure MPC_t. The key point is that this measure captures news about the expected medium-term policy path as opposed to news only about the contemporaneous policy decision, meaning that it encompasses the so-called target and path factors (Gürkaynak, Sack, and Swanson 2005a) of monetary news. Since there is no single measure of the overall stance of monetary policy during unconventional times, another advantage of this simple measure is that it can capture the multidimensional aspects of monetary policy such as extended liquidity provisions, forward guidance, or asset purchases which are also likely to affect interest rates at this horizon.[13]

We consider the surprise component of the IR publication as a reasonable proxy for surprises to central bank inflation and output projections that would enter in the central bank reaction function in standard macroeconomic models. We use the same measure as monetary news to capture IR news and compute the daily change in the two-year nominal gilt yield on IR publication dates (IR_t). Figure 1, shown later in this paper, plots the MPC and IR surprises over our sample. A simple visual inspection confirms the effect of the 2008–09

[13]Bean and Jenkinson (2001) suggest that the BoE is more likely to change policy in IR months, which would affect policy expectations. Our sample includes seven interest rate changes in IR months and eight in non-IR months.

financial crisis on the policy and macroeconomic outlook with large negative changes in both series around these dates.

With respect to the dependent variables considered, since the remit of the Bank of England's MPC is to target inflation, a natural candidate to investigate the effect of monetary policy is to measure their impact on inflation expectations. At the daily frequency, inflation swaps are a standard proxy for measuring compensation for expected inflation and inflation risk (Beechey, Johannsen, and Levin 2011).[14] These instruments are financial market contracts to transfer inflation risk from one counterparty to another. Most of the liquidity is driven by corporate firms at shorter maturities and pension, insurance, and retirement funds at longer maturities for hedging inflation exposures. We consider instantaneous forwards at different maturities, from two to five years ahead, which provide a proxy measure for expected inflation at the date of the maturity of the contract.[15] These are available since October 2004, which determines the starting date of our sample.[16] For robustness and comparison with the literature, we also consider daily returns in the Financial Times Stock Exchange (FTSE) price index, the share index of the 100 companies listed on the London Stock Exchange with the highest market capitalization.

We are then able to assess how private agents interpret monetary and IR surprises. We test the null hypothesis that tighter-than-expected monetary policy reduces inflation swaps, i.e., policy signals dominate macro signals. We also test the hypothesis that the publication of a higher economic prospects—a positive IR news surprise—increases inflation swaps, i.e., macro signals dominate policy signals. In equation (3), the two null hypotheses imply β_1 being negative and β_2 being positive. Table 1 presents our results for equation (3) estimated by ordinary least squares (OLS) using daily data.

---

[14]One advantage of these financial instruments is that they are directly related to payoff decisions. One drawback, however, is that they may be affected by term, liquidity, and inflation risk premiums.

[15]In the United Kingdom, they are linked to the Retail Price Index (RPI) measure of inflation, rather than the Consumer Prices Index (CPI), which is the measure the Bank's inflation target is currently based on.

[16]Table A.1 in the online appendix provides data sources and description while table A.2 provides some descriptive statistics.

Table 1. Estimates at the Daily Frequency

	(1) FTSE	(2) swap_2y	(3) swap_3y	(4) swap_4y	(5) swap_5y
MPC Surprises	0.111***	0.231*	0.256	0.199	0.156**
	(0.03)	(0.14)	(0.18)	(0.12)	(0.08)
IR Surprises	0.030	0.603**	0.423***	0.209***	0.041
	(0.02)	(0.25)	(0.11)	(0.07)	(0.11)
N	173	173	173	173	173
R^2	0.13	0.13	0.11	0.06	0.02
Separate Estimations					
MPC Surprises Only					
MPC Surprises	0.112***	0.233*	0.263	0.207*	0.162**
	(0.03)	(0.14)	(0.18)	(0.12)	(0.08)
N	130	130	130	130	130
R^2	0.15	0.03	0.04	0.03	0.03
IR Surprises Only					
IR Surprises	0.025	0.596**	0.394***	0.177**	0.012
	(0.02)	(0.25)	(0.11)	(0.07)	(0.11)
N	43	43	43	43	43
R^2	0.04	0.26	0.39	0.17	0.00

Notes: Heteroskedasticity-robust standard errors are in parentheses. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Each column corresponds to the OLS estimation of equation (3). The constant is equal to zero and never significant, so it has been removed from each panel for the sake of parsimony. The sample period goes from October 2004 to July 2015. The independent variables are the surprise component of MPC announcements and the surprise component of the IR publication, both computed as the daily change in two-year gilt nominal yields. The dependent variable is the daily change in FTSE and in inflation swaps at different maturities from two years to five years. The lower two panels show estimates when equation (3) is estimated on MPC or IR dates exclusively.

We compute heteroskedasticity-robust standard errors. Our sample period goes from October 2004 to July 2015 and covers 130 MPC announcements and the publication of 43 IRs. The independent variables are the surprise component of MPC announcements and the surprise component of the IR publication, both computed as the daily change in two-year nominal gilt yields. The dependent variables are the daily changes in inflation swaps at different maturities from two years to five years and stock prices.

Our results show that, first, the parameter associated with monetary surprises, β_1, is positive and significant at the 1 percent level for stock prices. It is also positive for inflation swaps at two- to five-year maturities, although it is only significant at the 5 percent level for the five-year maturity. This positive effect means that monetary surprises convey macro signals to private agents, in contrast to the prediction of full-information models and our null hypothesis, and consistent with Melosi (2017). Second, IR surprises are estimated to have no effect on stock prices and five-year-ahead inflation swaps, but they have a strong positive effect on inflation swaps at two- to four-year maturities. This suggests that IR surprises convey macro signals that dominate policy signals, consistent with our null hypothesis. We also estimate two separate equations for each type of surprise. The estimated parameters are very similar. It is interesting to note that IR surprises are estimated to explain 39 percent of the variance of inflation swaps three years ahead, while monetary surprises only explain 4 percent. In contrast, monetary surprises explain 15 percent of the variance of stock prices compared with 4 percent only for IR surprises. This difference suggests that monetary surprises and IR surprises, and their respective information content, are interpreted differently by private agents.

3.2 Using Co-movement to Disentangle Policy and Macro Signals

In the previous section, we have examined the overall effects of monetary and IR surprises. We now aim to provide intuition for the relevance of each signal embedded in these monetary and IR surprises. Following Cieslak and Schrimpf (2019) and Jarocinski and Karadi (2020), we use how policy and macro signals are expected to affect the co-movement of stock returns and interest rates following central bank policy announcements or the publication of central bank macroeconomic information.

Table 2. Identification through Co-movements

Signals	**Interest Rates**	**Stock Prices**	**Co-movement**
Policy	+	−	−
Macro	+	+	+

Policy signals are akin to monetary shocks and operate through the real short-term interest rate. A monetary tightening depresses stock prices by increasing the discount rate but increases longer-term yields through the expectations hypothesis of the term structure. The propagation mechanism of monetary policy across the term structure is confirmed by a number of studies (Rigobon and Sack 2004; Gürkaynak, Sack, and Swanson 2005a). Macro signals are similar to expected growth shocks and correspond to private expectations about real activity. In the Gordon growth model, positive shocks to growth expectations raise stock prices via the cash-flow news channel. Such shocks also lead to an increase in interest rates, through the investment channel and the response of policymakers to real activity news. The effect of real activity news on the yield curve is typically found to be positive and hump shaped (e.g., Gürkaynak, Sack, and Swanson 2005b). Table 2 summarizes the expected effects of policy and macro signals on stock returns and the yield curve, and their co-movement.

Both interest rates and stock prices co-vary negatively in response to policy signals, akin to monetary policy shocks, but both co-vary positively in response to macro signals, akin to growth shocks. Therefore, in order to disentangle the two signals, we compute the co-movement of daily changes in interest rates and stock prices, $\rho(r,s)$. We are then left with two options. The first option is to consider monetary (or IR) surprises as mainly about policy signals when the co-movement is negative and mainly about macro signals when the co-movement is positive. This discrete categorization has the benefit of simplicity but may hide some subtlety in the way the signals are perceived by private agents. In order to avoid to consider all positive (resp. negative) co-movement as a macro (policy) signal, we estimate a transition function such that we can decompose monetary (or IR) surprises in two components based on the strength of the co-movement. Monetary (or IR) surprises are then represented as weighted average of the two signals, the weight being obtained from the transition function. We compute the following transition function:

$$\mathrm{F}(\rho_t) = 1 - (\exp(-\gamma\rho_t)/(1 + \exp(-\gamma\rho_t))). \tag{4}$$

F(.) is a smooth transition function with a range between 0 (negative co-movement) and 1 (positive co-movement), ρ_t is a variable measuring the co-movement, and γ is the parameter governing the degree of smoothness of the transition from one state to the other with $\gamma > 0$. We can thus decompose daily monetary (or IR) surprises as a linear combination of policy and macro signals, $\mathrm{MPC}_t = \mathrm{S}_t^{\mathrm{pol}} + \mathrm{S}_t^{\mathrm{macro}}$, using the weight $\mathrm{F}(\rho_t)$ such that $\mathrm{S}_t^{\mathrm{pol}} = \mathrm{MPC}_t \times (1 - \mathrm{F}(\rho_t))$ and $\mathrm{S}_t^{\mathrm{macro}} = \mathrm{MPC}_t \times (\mathrm{F}(\rho_t))$.

Using the co-movement in stock prices and interest rates and the estimated weight, we compute policy and macro signals for both monetary and IR surprises. Table 3 presents our results for equation (3) modified to include the signal decomposition for each of the two types of surprises, and estimated by OLS using daily data. Our sample period is unchanged and covers 130 MPC announcements and the publication of 43 IRs. The independent variables are the surprise component of MPC announcements and the surprise component of the IR publication, both computed as the daily change in two-year gilt nominal yields. The dependent variable is the daily change in stock prices and inflation swaps at different maturities from two years to five years. We compute heteroskedasticity-robust standard errors.

Looking at the responses of FTSE returns, the parameters associated with policy signals, whether they come from monetary or IR surprises, have a negative impact whereas macro signals have a positive impact. This is also true for inflation swaps, with the effect of macro signals from monetary surprises dominating the effect of policy signals, consistent with table 1. Regarding the effect of IR surprises on inflation swaps, the policy signals conveyed have no effect (except at the one-year maturity) whereas macro signals have a positive impact. These outcomes are confirmed when estimating two separate equations for each type of surprise. These findings suggest first that policy and macro signals affect inflation swaps

Table 3. Co-movement Identification of Policy and Macro Signals

	(1) FTSE	(2) swap_2y	(3) swap_3y	(4) swap_4y	(5) swap_5y
MPC Policy Signals	−0.435***	−0.762**	−1.791***	−1.078***	−0.254
	(0.13)	(0.36)	(0.38)	(0.24)	(0.18)
MPC Macro Signals	0.522***	0.985***	1.798***	1.161***	0.464***
	(0.12)	(0.36)	(0.27)	(0.14)	(0.10)
IR Policy Signals	−0.254***	−1.760**	−0.373	−0.076	−0.134
	(0.07)	(0.76)	(0.44)	(0.22)	(0.50)
IR Macro Signals	0.277***	2.655***	1.119***	0.461**	0.194
	(0.07)	(0.65)	(0.37)	(0.19)	(0.30)
N	173	173	173	173	173
R^2	0.37	0.26	0.23	0.13	0.03
Separate Estimations					
MPC Surprises Only					
MPC Policy Signals	−0.435***	−0.762**	1.793***	−1.080**	−0.256
	(0.13)	(0.35)	(0.38)	(0.24)	(0.18)
MPC Macro Signals	0.524***	0.982***	1.810***	1.176***	0.478***
	(0.12)	(0.36)	(0.27)	(0.14)	(0.10)
N	130	130	130	130	130
R^2	0.36	0.06	0.17	0.12	0.04
IR Surprises Only					
IR Policy Signals	−0.253***	−1.760**	−0.373	−0.076	−0.133
	(0.08)	(0.78)	(0.45)	(0.22)	(0.51)
IR Macro Signals	0.270***	2.668***	1.068***	0.400**	0.141
	(0.07)	(0.67)	(0.38)	(0.20)	(0.31)
N	43	43	43	43	43
R^2	0.42	0.55	0.49	0.20	0.01

Notes: Heteroskedasticity-robust standard errors are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Each column corresponds to the OLS estimation of equation (3) where MPC and IR surprises are replaced by their respective policy and macro signals based on equation (4). The constant is equal to zero and never significant, so it has been removed from each panel for the sake of parsimony. The sample period goes from October 2004 to July 2015. The independent variables are the surprise component of MPC announcements and the surprise component of the IR publication, both computed as the daily change in two-year gilt nominal yields. The dependent variable is the daily change in FTSE and in inflation swaps at different maturities from two years to five years. The lower two panels show estimates when equation (3) is estimated on MPC or IR dates exclusively.

in the expected direction and second that MPC policy decisions have a strong macroeconomic content consistent with the signaling channel of Melosi (2017). He finds that inflation expectations may respond positively to contractionary monetary shocks under certain calibrated parameters. If the quality of private information is poor relative to that of central bank information, and/or if the policy rate is more informative about nonmonetary shocks than about monetary shocks, then the macro outlook signaling channel may be at work. Similarly, Tang (2015) finds such a positive effect when prior uncertainty about inflation is high. At the opposite, the publication of the IR has almost no policy content, i.e., private agents do not interpret IR surprises as news about future policy, but only about the future state of the economy.

A potential concern with the specification of equation (3) is that the effect of monetary and IR surprises could reflect other macroeconomic news published at these dates. To address this concern, we augment equation (3) with news surprises in six of the most important macroeconomic data releases that have been released on MPC or IR dates: employment change, International Labour Organisation (ILO) unemployment rate, industrial production, PMI Services, average weekly earnings, and producer price index (PPI). These surprises are computed as the difference between actual releases and Bloomberg surveys on the days before the release. For instance, industrial production has been released 30 times on MPC dates over our sample, while earnings and unemployment have been published 27 times each on IR dates.

Table A.3 in the online appendix (available at http://www.ijcb.org) presents the results for equation (3) augmented with these news surprises. The sign and magnitude of the effect both policy and macro signals of monetary and IR surprises is confirmed. Table A.3 also includes a specification in which surprises are computed as change in the one-year interest rate, following the maturity used in Gürkaynak, Sack, and Swanson (2005a), and a specification based on monetary surprises from Cesa-Bianchi, Thwaites, and Vicondoa (2020) obtained with intraday data and a 30-minute window around policy announcements. Finally, equation (3) is estimated on a subsample starting on September 15, 2008 with the bankruptcy of Lehman Brothers. These estimates confirm the respective dominant signal of monetary and IR surprises.

4. The Effect of Monetary and Projection Shocks at the Monthly Frequency

Since policy decisions happen every month and central bank projections are published every quarter, an additional way to assess the effects of policy and macro signals is to work at a monthly frequency. Doing so has at least three advantages. First, it is possible to estimate monetary shocks—i.e., shocks to the policy instrument—in contrast with monetary surprises that are akin to shocks to private agents' information set. Identifying monetary shocks enables us to control for changes in policymakers' information set. Second, we are able to identify central bank projection surprises.[17] Third, it enables us to use a standard inflation expectation formation model derived from the information friction literature.

4.1 Estimating Exogenous Innovations

Estimating the effects of the policy rate and Bank's inflation and output projections raises a major econometric challenge mentioned earlier: at the monthly frequency, our three variables of interest are likely to be endogenous to inflation expectations. We therefore perform first-stage regressions to extract the unpredictable components of $i, \pi^{\mathrm{CB}}, x^{\mathrm{CB}}$ orthogonal to their systematic components. We follow the Romer and Romer (2004) approach applied to U.K. data by Cloyne and Huertgen (2016) such that we remove the contribution of the most relevant endogenous factors that would underlie the evolution of these variables.

In order to cope with the potential presence of different information sets (see Blanchard, L'Huillier, and Lorenzoni 2013, and Miranda-Agrippino and Ricco 2020), we make sure that exogenous innovations are orthogonal not only to the central bank's information set but also to private agents' information set. The inclusion of both private and central bank forecasts in the regression model enables us to deal with three concerns. First, forecasts encompass rich information sets and work as a factor-augmented vector autoregressive

[17] Months during which no central bank projections are published are treated as missing observations, since no new information is conveyed by policymakers. We detail how we proceed more specifically later on.

(FAVAR) model (see Bernanke, Boivin, and Eliasz 2005). Second, forecasts are not revised so they capture real-time information. Third, forecasts capture forward-looking information sets.

Our identification works through timing restrictions. We aim to remove the contribution of *lagged* macro and private forecasts (so that innovations can have contemporaneous effects on these) and the contribution of *contemporaneous* Bank variables (to remove the information of policymakers). Starting with the monetary shock, we estimate the following equation:

$$i_t = f\left(\pi^{CB}_{t,pca}, x^{CB}_{t,pca}, mc_{t,pca}, \Psi_{t-1}, I_{t,IR}\right) + \varepsilon^i_t. \tag{5}$$

We assume that changes in i_t are driven by the policymakers' response to its own inflation $\pi^{\mathrm{CB}}_{t,pca}$ and output $x^{\mathrm{CB}}_{t,pca}$ projections; to the market interest rate curve ($mc_{t,pca}$) used as conditioning path for the Bank's macroeconomic projections; to a vector (Ψ_{t-1}) comprising private inflation, output, and short-term interest rate expectations as well as macroeconomic controls; and to whether the policy decision is made during IR months ($I_{t,IR}$) or not. $f(\cdot)$ is the function capturing its systematic reaction, and the error term ε^i_t reflects monetary shocks orthogonal to the macroeconomic outlook.

We introduce inflation and output projections using the respective first principal component of the Bank's inflation and output projections for different maturities. We also include $mc_{t,pca}$, the first principal component of the market interest rate curve at the one- to three-year maturities used as conditioning path for the Bank's macroeconomic projections. The vector Ψ_{t-1} includes a lag of the change in the policy rate capturing the last policy decision, a lag of the first principal component of inflation swaps from 1 to 10 years ahead, a lag of the first principal component of private output forecasts from 1 quarter to 3 years ahead measured from surveys, and a lag of the first principal component of private short-term interest rate forecasts from 1 to 10 years ahead measured from nominal government bonds.[18] The vector Ψ_{t-1} also includes the following macroeconomic controls that are likely to affect future

[18]We use a principal component analysis and consider the first principal component for each forecast variable so as not to include all horizons in the estimated model and avoid multicollinearity. The first principal component intends to capture the overall forward-looking information set of forecasters for all horizons together. The first principal component of inflation swaps captures 76 percent of

inflation dynamics and hence inflation expectations: CPI inflation, industrial production, oil prices, net lending, the sterling effective exchange rate, and housing prices (all included as annual growth rates) together with dummies for when the forward-guidance policy started and when the economy has reached its ELB. These macroeconomic controls are grouped in a vector denoted Z_t.[19] Finally, we also include a dummy for when the Bank publishes its Inflation Report (IR). It is meant to capture that expectations of policy decisions may be different in IR and non-IR months, as private agents may expect the central bank to update its policy more frequently during IR months when it updates its assessment of the state of the economy.[20]

During a significant part of the sample period considered, the policy rate has been at the ELB, and monetary policy has taken various dimensions in the meantime. In a complementary specification, the policy instrument is proxied by a shadow rate that measures U.K. monetary policy as the U.K. policy rate (Bank Rate) until 2009 and then mechanically adjusts for the quantitative easing (QE) undertaken by the Bank of England's Monetary Policy Committee to capture the overall stance of monetary policy.[21]

Central bank inflation or output projection shocks at a given horizon should be seen as the unpredictable innovation of a projection at a given horizon, conditional on the information available to private agents at the date when the projections are published.

---

the common variance of the underlying series, the one of the market curve 97 percent, and the one of private output forecasts 85 percent. In the robustness section, we replace these first principal components with the underlying series at all different horizons.

[19]Some of these macroeconomic controls are revised vintages. We acknowledge that this is a limitation of the present study since due to data availability we cannot work with true real-time data.

[20]While Bean and Jenkinson (2001) report that the Bank is more likely to change interest rates in Inflation Report months, our sample includes seven interest rate changes in IR months and eight changes in non-IR months.

[21]As in Forbes, Hjortsoe, and Nenova (2018), the shadow rate series is mechanically constructed by comparing the estimated effects of QE to the economic multipliers assigned to conventional changes in Bank Rate. For further detail on the economic impact of U.K. asset purchases, see Joyce, Tong, and Woods (2011). The underlying assumption is that QE is a close substitute as a monetary policy instrument to Bank Rate such that the lower bound was not an effective constraint on monetary policy over the period in question.

We estimate these shocks by using the Bank's inflation and output projections conditioned on the path for Bank Rate implied by market interest rates prior to the policy meeting, so independent from the policy decision. We estimate the following two equations, for inflation and output projections, respectively:

$$\pi_{t,h}^{CB} = g(mc_{t,pca}, \Psi_{t-1}', I_{t,ZLB}) + \varepsilon_t^{\pi_{CB,h}} \tag{6}$$

$$x_{t,h}^{\text{CB}} = g'(mc_{t,pca}, \Psi_{t-1}', I_{t,ZLB}) + \varepsilon_t^{x_{CB,h}}. \tag{7}$$

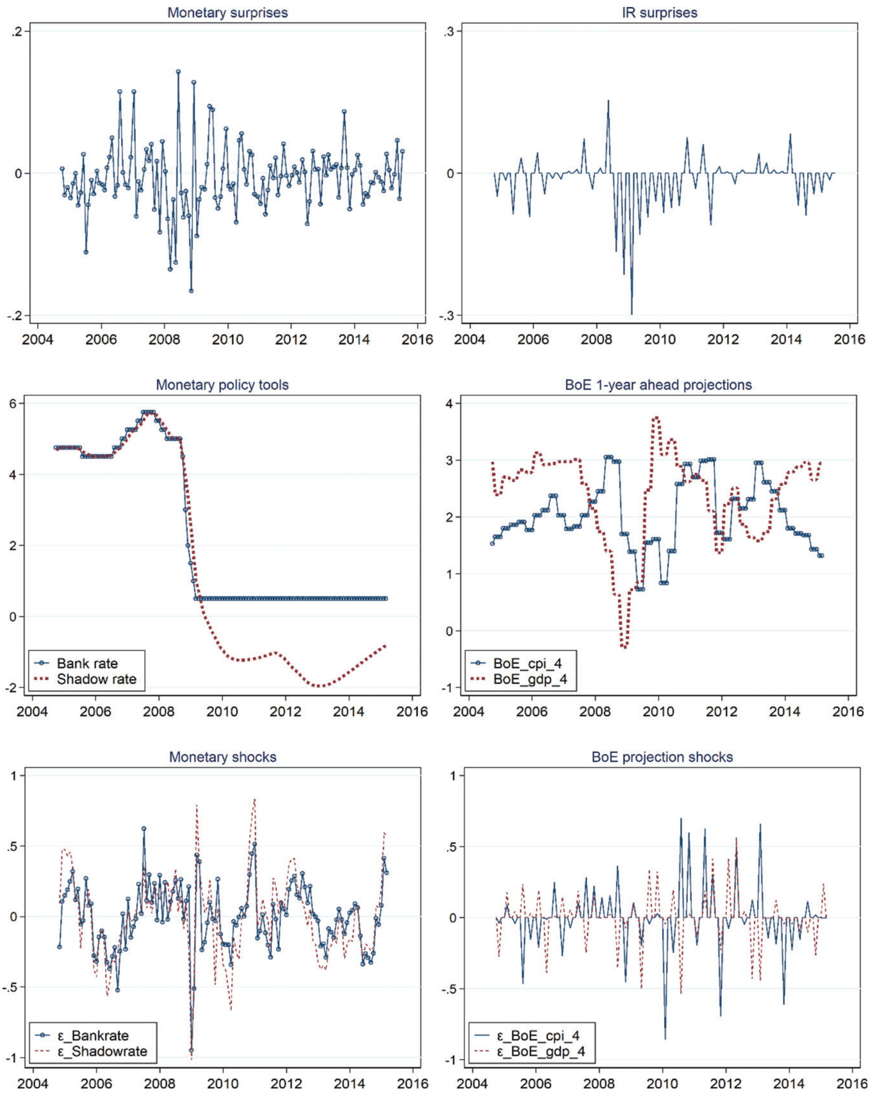
So changes in $\pi_{t,h}^{\text{CB}}$ and $x_{t,h}^{\text{CB}}$ are driven by $mc_{t,pca}$, the first principal component of the market interest rate curve at the one- to three-year maturities. The vector Ψ_{t-1}' includes a lag of the first principal components of the Bank's inflation and output projections, a lag of the policy (or shadow) rate, a lag of the first principal component of private inflation expectations, private output forecasts, and private short-term interest rate forecasts, together with the vector of macroeconomic controls Z_{t-1}. We also add a dummy for when Bank Rate is at its effective lower bound. Equations (6) and (7) are estimated on IR months only, since no projections are published during non-IR months (during which, by construction, projection shocks are zero).[22] This is performed without affecting the lag structure—for instance, February projections take January values for the lagged variables.

Figure 1 plots the estimated monetary and projection shocks. Table A.4 in the online appendix shows the estimated parameters of equations (7)–(9), and the properties and the correlation structure of innovations. For these estimated exogenous series to be relevant, they should be unpredictable from movements in data. We assess the predictability of these series with Granger-causality type tests.[23] The F-stats and adjusted R^2 shown in table A.4 suggest that the

[22] An alternative, tested in the robustness section, is to proceed to a constant-interpolation of projection shocks for the following two months after the IR publication, as one could argue that these projection shocks are available to private agents updating their information set later on.

[23] We consider a set of standard macro and financial variables: inflation, industrial production, oil prices, wages, net lending, the U.K. move, and expected future short-term interest rates.

Figure 1. Daily-Frequency Monetary and IR Surprises and Monthly-Frequency Monetary and Projection Shocks



Notes: The first row shows MPC and IR surprises. The surprise component of MPC decisions and of the IR publication are both computed as the daily change in two-year gilt nominal yields on MPC and IR dates. The second row shows the value of Bank Rate and the Bank of England's shadow rate together with inflation and output projections one year ahead. The third row shows exogenous shocks to these variables estimated from equations (5) through (7). Parameters are presented in table A.4 in the online appendix.

null hypothesis that these estimated innovations are unpredictable cannot be rejected.

4.2 The Empirical Model

Our empirical setup is motivated by two theoretical models with rational expectations and information frictions. In the sticky-information model of Mankiw and Reis (2002) and Carroll (2003), private agents update their information set infrequently as they face costs of absorbing and processing information. When private agents update their information set, they gain perfect information. In the noisy-information models of Woodford (2001) and Sims (2003), private agents continuously update their information set but observe only noisy signals about the true state of the economy. Their inertial reaction arises from the inability to pay attention to all the information available. Internalizing their information-processing capacity constraint, they remain inattentive to a part of the information (Moscarini 2004).

Under the assumption that private agents have homogeneous inflation expectations, we can bridge these two strands of the literature in a simple and general specification.[24] Private inflation expectations—the average of daily observations in each month—are modeled as a linear combination of prior beliefs about future inflation, lagged expectations $\pi^{PF}_{t-1,h}$, and new (and potentially noisy) information relevant for future inflation released between $t-1$ and t, measured by the vector Λ_t:[25,26]

[24]This assumption matches the point forecasts nature of inflation swaps. We acknowledge that point forecasts may suffer an aggregation bias because agents may have heterogeneous beliefs due to differences in their own information sets, but we abstract from this issue in this paper.

[25]Section A.1 of the online appendix describes the regression-based approach used to correct inflation swaps in levels for potential term, liquidity, and inflation risk premiums, and extract inflation expectations.

[26]In sticky- and noisy-information models, expectations for a given date in the future can be represented as a weighted average of new information and previous expectations for the same date in the future. Because of data limitations, it is not possible to test exactly this specification. We therefore need to assume that the private forecast process is persistent to bridge our empirical exercise to these models. This has been shown by Coibion and Gorodnichenko (2012, 2015) and Andrade and Le Bihan (2013), among others.

$$\pi^{PF}_{t,h} = \beta_0 + \beta_L \pi^{PF}_{t-1,h} + \beta_\Lambda \Lambda_t + \varepsilon_t. \tag{8}$$

This specification allows us to be agnostic about the nature of information frictions.[27] The vector Λ_t could include any variable that is likely to affect inflation and that can be used to predict future inflation. We decompose this vector into three groups of variables.

The first group comprises our externally identified instruments for monetary shocks (ε^i_t) as well as inflation ($\varepsilon^{\pi CB,h}_t$) and output ($\varepsilon^{xCB,h}_t$) projection shocks. To test our research question, we explicitly assume that these surprises are incorporated in private agents' forecasting function.[28] A second group includes the macroeconomic variables listed in the vector Z_t. A third group, denoted with the vector X_t, aims to capture news shocks and surprises to macro developments that are contemporaneous to monetary and projection shocks. It includes a news variable π^s which captures the information content of any data released between $t-1$ and t that may affect inflation (see Andersen et al. 2003). This news variable is defined as the difference between the actual value of CPI inflation in t and private inflation forecasts formed at date $t-1$ for the quarter t ($\pi^s = \pi_t - \mathrm{E}_{t-1}\pi_t$). This is equivalent to an inflation forecast error. Bloomberg Consensus provides the market average expected one-month-ahead CPI inflation at a monthly frequency. X_t also comprises the change between $t-1$ and t in private output and interest rate forecasts, to control for their link with private inflation forecasts as evidenced by Fendel, Lis, and Rülke (2011), Paloviita and Viren (2013), and Dräger, Lamla, and Pfajfar (2016). Finally, we capture the presence of macro news by using the three indexes estimated by Scotti (2016) for the United Kingdom: the real activity index, capturing the state of economic conditions; the surprise index,

[27]The value of β_L, expected to be positive, sheds light on whether the limited adjustment mechanism in which information is only partially absorbed over time is at work in the data.

[28]The timing of policy decisions and IR releases—detailed in section 2—which are made public in the early days of the given months should ensure that their information content is not already contained in private inflation expectations and that inflation expectation dynamics are not responsible for these shocks. We test the robustness of this assumption by considering only the last daily observation of each month for our left-hand-side variable so as to remove any potential endogeneity issue.

summarizing economic data surprises; and the uncertainty index, measuring uncertainty related to the state of the economy.

Equation (8) can be written as

$$\pi^{PF}_{t,h} = \beta_0 + \beta_L \pi^{PF}_{t-1,h} + \beta_1 \varepsilon^i_t + \beta_2 \varepsilon^{\pi CB,h}_t + \beta_3 \varepsilon^{xCB,h}_t + \beta_Z Z_t + \beta_X X_t + \varepsilon_t, \quad (9)$$

where ε^i_t, $\varepsilon^{\pi CB,h}_t$, and $\varepsilon^{xCB,h}_t$ are the monetary and projection shocks—h being the horizon of the given projection. The sign of the $\beta_1 - \beta_3$ parameters sheds light on how monetary shocks and inflation or output projection shocks are interpreted by private agents. If the macro signal dominates, these parameters will be positive; if the policy signal dominates, they will be negative.

4.3 The Effect of Monetary and Projection Shocks

We estimate equation (9) with OLS for the term structure of inflation expectations and stock prices and compute heteroskedasticity-robust standard errors.[29] Our baseline analysis focuses on central bank projections four quarters ahead. This horizon falls before interest rates are generally estimated to have their peak effect on inflation—around 18–24 months ahead—and therefore enables us to minimize the control issue.[30] On the other hand, the shortest horizon of the term structure of inflation expectations studied here (two years) falls after this horizon of central bank projections, so strategic forecasting motives should be absent.

Table 4 provides evidence that contractionary shocks to Bank Rate decrease stock prices and private inflation expectations at all horizons from two to five years ahead—β_1 is negative. That is consistent with contractionary policy shocks affecting these variables through the usual transmission mechanism channel and suggests that a policy signal is taken from monetary shocks. The magnitude of the effect decreases with the horizon of expectations, consistent

[29]Because our empirical strategy proceeds in two steps, we also estimate equation (9) with bootstrapped standard errors so we can take into account the estimation uncertainty from the first-stage regressions (equations (5)–(7)).

[30]The interest rate instrument gives the central bank some control over the forecasted variables, and this issue is circumvented when the horizon of forecasts is shorter than the transmission lag of monetary policy.

Table 4. Monthly-Frequency Estimates

	(1) FTSE	**(2) PF_2y**	**(3) PF_3y**	**(4) PF_4y**	**(5) PF_5y**
A. Baseline					
ε_Bankrate	−0.033*	−0.204**	−0.171**	−0.154**	−0.140**
	(0.02)	(0.09)	(0.08)	(0.07)	(0.07)
ε_BoE_cpi_4	−0.022	0.217**	0.181**	0.138*	0.099*
	(0.03)	(0.11)	(0.09)	(0.07)	(0.06)
ε_BoE_gdp_4	0.004	0.096	0.074	0.042	0.007
	(0.03)	(0.11)	(0.09)	(0.08)	(0.08)
Lag. Dep. Var.	0.672***	0.728***	0.767***	0.793***	0.807***
	(0.07)	(0.09)	(0.09)	(0.08)	(0.08)
Constant	0.096*	0.662**	0.555**	0.529**	0.533**
	(0.05)	(0.26)	(0.24)	(0.23)	(0.22)
Controls: X_t and Z_t	Yes	Yes	Yes	Yes	Yes
N	125	125	125	125	125
R^2	0.91	0.70	0.75	0.78	0.81
B. Separate Estimations					
Bank Rate Only					
ε_Bankrate	−0.032*	−0.189**	−0.158**	−0.147**	−0.138**
	(0.02)	(0.09)	(0.08)	(0.07)	(0.06)
CB Projections Only					
ε_BoE_cpi_4	−0.022	0.214*	0.178*	0.136*	0.097*
	(0.03)	(0.11)	(0.09)	(0.07)	(0.06)
ε_BoE_gdp_4	−0.007	0.031	0.019	−0.007	−0.038
	(0.03)	(0.10)	(0.09)	(0.08)	(0.07)
C. Bootstrapped Standard Errors					
ε_Bankrate	−0.033	−0.204*	−0.171*	−0.154*	−0.140*
	(0.02)	(0.11)	(0.09)	(0.08)	(0.07)
ε_BoE_cpi_4	−0.022	0.217*	0.181*	0.138*	0.099
	(0.03)	(0.12)	(0.10)	(0.08)	(0.06)
ε_BoE_gdp_4	0.004	0.096	0.074	0.042	0.007
	(0.04)	(0.12)	(0.10)	(0.09)	(0.08)
D. Shadow Rate					
ε_Shadowrate	−0.056***	−0.151**	−0.130**	−0.114**	−0.100*
	(0.02)	(0.07)	(0.06)	(0.05)	(0.05)
ε_BoE_cpi_4	−0.021	0.222**	0.186**	0.143**	0.103*
	(0.03)	(0.11)	(0.09)	(0.07)	(0.06)
ε_BoE_gdp_4	0.016	0.086	0.067	0.035	−0.001
	(0.03)	(0.10)	(0.08)	(0.08)	(0.07)

(*continued*)

Table 4. (Continued)

	(1) FTSE	(2) PF_2y	(3) PF_3y	(4) PF_4y	(5) PF_5y
Lag. Dep. Var.	0.611***	0.738***	0.775***	0.797***	0.807***
	(0.08)	(0.10)	(0.09)	(0.09)	(0.08)
Constant	0.101*	0.660**	0.553**	0.532**	0.541**
	(0.05)	(0.27)	(0.25)	(0.24)	(0.23)
Controls: X_t and Z_t	Yes	Yes	Yes	Yes	Yes
N	125	125	125	125	125
R^2	0.92	0.70	0.75	0.78	0.81

Notes: Heteroskedasticity-robust standard errors are in parentheses. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Each column corresponds to the OLS estimation of equation (9). The sample period goes from October 2004 to July 2015. Monetary and projection shocks are estimated based on equations (5)–(7). The dependent variable is the level of monthly averaged FTSE price index and inflation swaps at different maturities from two years to five years. These swaps are corrected for premiums based on equation (A.2). For parsimony, only the key coefficients are reported. Complete tables are available from the authors upon request. X_t includes a news variable capturing the information flow between $t-1$ and t of macro data releases related to inflation, the change between $t-1$ and t in private output and interest rate forecasts, the real activity, uncertainty, and news indexes of Scotti (2016). Z_t includes CPI, industrial production, oil prices, the sterling effective exchange rate, net lending, housing prices, and forward-guidance (FG) and zero lower bound (ZLB) dummies. Panel B shows estimates when equation (9) is estimated with monetary shocks or projection shocks separately. Panel C shows estimates with bootstrapped standard errors to take into account the estimation uncertainty from equations (5)–(7). Panel D shows estimates when equation (9) is estimated with monetary shocks based on a shadow rate measure.

with waning effects of monetary policy on inflation. The transmission lags of monetary policy are often estimated to be around 18 to 24 months for inflation (see, e.g., Bernanke and Blinder 1992 or Bernanke and Mihov 1998), so negative effects at horizons longer than the transmission lags could be interpreted as a policy signal effect going through the expectations channel.[31] Even if macro signals may still be at work, there is no evidence that it outweighs the policy signal given the consistently negative effect.

[31]Fatum and Hutchison (1999) find no evidence in the United States supporting the policy signaling hypothesis that policy actions are related to changes in expectations about the stance of future monetary policy. However, their analysis focuses specifically on foreign exchange market interventions.

We also test whether the dominant signal that the Bank's inflation and output projections convey is about the state of the economy or about the policy path. Table 4 reports that positive shocks to the Bank's inflation projections at four quarters ahead have no significant effect on stock prices but increase private inflation expectations two to five years ahead—β_2 is positive. That effect is strongest two years ahead and again decreases with the horizon. The sign of the effect suggests that the information conveyed about the macro outlook outweighs the policy signal conveyed by these projections. That is consistent with private agents and the central bank having different information sets. At the opposite, shocks to the Bank's output projections have no effect.

Table 4 also shows estimates when Bank Rate is replaced by a shadow rate measure in order to take into account the ELB and the implementation of various unconventional measures. The results are similar: contractionary monetary shocks have a negative effect on inflation expectations and stock prices. In order to check that the signals conveyed by each piece of information are specific to it (i.e., that an inflation projection shock is not interpreted differently when we include or do not include monetary shocks in our specification, for instance), we have also estimated equation (9) with monetary shocks or projection surprises alone. These additional estimates, also shown in table 4, confirm the stability of the baseline results.

Table 5 shows that the Bank's inflation projections 8 quarters ahead also have a positive effect on inflation expectations between two and five years ahead, although those 12 quarters ahead have no effect. Output projections, at 8 or 12 quarters ahead, still have no effect. The fact that surprises to the Bank's short- and medium-term inflation projections affect private inflation expectations at medium- and long-term horizons suggests that private agents take a signal about the inflation outlook further ahead. This result holds when considering shocks to the shadow rate as well.

Overall, these results suggest that, in contrast to the theoretical predictions of full-information models described in section 2.1, there is some evidence that some weight is placed on the signals about the macro outlook that projections contain. One interpretation of these results is that when private agents face a signal-extraction problem from one piece of information, they rely on the underlying nature of the information disclosed by the central bank: a monetary shock

Table 5. Longer-Horizon BoE Projection Estimates

	(1) FTSE	(2) PF_2y	(3) PF_3y	(4) PF_4y	(5) PF_5y
8-Quarter-Ahead BoE Projections					
ε_Bankrate	−0.034*	−0.216**	−0.182**	−0.167**	−0.156**
	(0.02)	(0.09)	(0.08)	(0.07)	(0.07)
ε_BoE_cpi_8	0.026	0.404*	0.364**	0.313**	0.263**
	(0.06)	(0.22)	(0.18)	(0.15)	(0.13)
ε_BoE_gdp_8	0.02	0.184	0.154	0.137	0.118
	(0.03)	(0.15)	(0.12)	(0.10)	(0.08)
N	125	125	125	125	125
R^2	0.91	0.69	0.74	0.78	0.81
Shadow Rate					
ε_Shadowrate	−0.058***	−0.164**	−0.142**	−0.128**	−0.116**
	(0.02)	(0.07)	(0.06)	(0.05)	(0.05)
ε_BoE_cpi_8	0.048	0.432**	0.390**	0.336**	0.282**
	(0.06)	(0.21)	(0.17)	(0.15)	(0.12)
ε_BoE_gdp_8	0.034	0.185	0.156	0.138	0.118
	(0.03)	(0.15)	(0.12)	(0.09)	(0.08)
N	125	125	125	125	125
R^2	0.92	0.69	0.74	0.78	0.81
12-Quarter-Ahead BoE Projections					
ε_Bankrate	−0.036**	−0.189**	−0.161**	−0.151**	−0.144**
	(0.02)	(0.09)	(0.08)	(0.07)	(0.07)
ε_BoE_cpi_12	0.109	0.078	0.094	0.093	0.084
	(0.08)	(0.22)	(0.18)	(0.16)	(0.15)
ε_BoE_gdp_12	0.012	−0.042	−0.015	0.022	0.053
	(0.05)	(0.14)	(0.12)	(0.11)	(0.10)
N	125	125	125	125	125
R^2	0.91	0.68	0.73	0.77	0.81
Shadow Rate					
ε_Shadowrate	−0.058***	−0.138**	−0.121**	−0.111**	−0.103**
	(0.02)	(0.07)	(0.06)	(0.05)	(0.05)
ε_BoE_cpi_8	0.131*	0.086	0.103	0.101	0.089
	(0.07)	(0.23)	(0.18)	(0.16)	(0.15)
ε_BoE_gdp_8	0.016	−0.047	−0.018	0.019	0.049
	(0.04)	(0.14)	(0.12)	(0.11)	(0.10)

(*continued*)

Table 5. (Continued)

	(1) FTSE	(2) PF_2y	(3) PF_3y	(4) PF_4y	(5) PF_5y
N	125	125	125	125	125
R^2	0.92	0.68	0.73	0.77	0.80

Notes: Heteroskedasticity-robust standard errors are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Each column corresponds to the OLS estimation of equation (9). The sample period goes from October 2004 to July 2015. Monetary and projection shocks are estimated based on equations (5)–(7). The dependent variable is the level of monthly averaged FTSE price index and inflation swaps at different maturities from two years to five years. These swaps are corrected for premiums based on equation (A.2). For parsimony, only the key coefficients are reported. Complete tables are available from the authors upon request. X_t includes a news variable capturing the information flow between $t-1$ and t of macro data releases related to inflation, the change between $t-1$ and t in private output and interest rate forecasts, the real activity, uncertainty, and news indexes of Scotti (2016). Z_t includes CPI, industrial production, oil prices, the sterling effective exchange rate, net lending, housing prices, and FG and ZLB dummies.

primarily conveys a policy signal and a projection shock primarily conveys a macro outlook signal.

4.4 Sensitivity Analysis

We run several alternative tests to assess the robustness of the baseline results. They are decomposed into tests about the identification of the monetary and projection shocks, and about additional right-hand-side variables and subsample estimates.

Focusing on the identification of monetary shocks, we first identify monetary shocks without controlling for the specificity of IR months. Second, we assess the effects of big and small monetary shocks (greater and lesser than 25 basis points) to control for potential outliers. Third, because the ELB may affect how private agents form their expectations, we estimate equation (5) on two subsamples pre- and post-ELB. Fourth, we replace the first principal components of private expectations in equations (5)–(7) with all individual series at different horizons. Table A.7 in the online appendix confirms the dominant signal for monetary and projection shocks, respectively.

We assess the impact of variations to projection shocks. First, we use a constant-interpolated measure of the projection shocks, so the observations during the two months after the IR publication take the value of the shock happening in the first month instead of zeros. Second, we interpolate projections the same way and estimate equations (6) and (7) on all observations rather than on IR months only. Third, we replace the shadow rate measure with the Bank Rate in equation (5). Table A.8 in the online appendix confirms the sign and magnitude of the main effects.

Because the variables in X_t and Z_t in equation (9) are also likely to be endogenous to inflation expectations, the estimation of the parameters associated to monetary and projection shocks may be biased. We estimate equation (9) with no controls to examine whether this potential bias affects our results. In addition, because contemporaneous news shocks may affect inflation expectations as well as central bank projections, the estimation would require controlling for as many news shocks as possible. First, we augment equation (9) with different measures of inflation: retail prices (RPI), core CPI, and input producer prices (PPI). Second, we augment equation (9) with different measures of economic slack: unemployment, capacity utilization, and the output gap computed by the Office for Budget Responsibility (OBR). Third, we augment equation (9) with different measures of financial stress: the Chicago Board Options Exchange Volatility Index (VIX), the U.K. move, and the St. Louis Fed Financial Stress Index, three daily-frequency indexes reacting in real time to macroeconomic and financial developments. Fourth, we augment equation (9) with a value added tax (VAT) dummy which takes the value of one in December 2008, January 2010, and January 2011 when the U.K. government raised the VAT, causing inflation to rise. Finally, we augment the model with the three European Commission (EC) U.K. sentiment measures for the industry, services, and consumers. Table A.9 in the online appendix confirms the magnitude and sign of the effects of monetary and projection shocks.

We estimate equation (9) on two subsamples before and after March 2009, when Bank Rate reached its ELB, so as to investigate the robustness of our results when Bank Rate was considered the main policy instrument and when macroeconomic dynamics may be affected by the policy rate being at the ELB. In addition, we test a specification in which we introduce a dummy for the dates of the announcements of explicit forward guidance on future policy rates in August 2013 and February 2014.[32] Table A.10 in the online appendix shows that our main results are robust to these alternatives.

We conclude this robustness section with another set of tests related to the specification of the left-hand-side variables. First, we replace the level of inflation expectations with their first difference. Second, we replace the level with their deviation from the Bank's inflation target (corrected for the sample mean of the wedge between RPI and CPI).[33] Third, we consider a more extreme information assumption, replacing the monthly average of all observations with the last daily observation of the month. While this option discards all data points before the last observation, it ensures that (i) all shocks or news released during a month are available to private agents and incorporated in the last observation; and (ii) there is no endogeneity issue between left-hand-side variables and explanatory variables. Table A.11 in the online appendix shows that the sign and magnitude of monetary and projection shocks are confirmed.

5. Conclusion

This paper investigates the effect of two types of central bank announcements private inflation expectations. We find that policy decisions convey some signals about the macroeconomic outlook, but that policy signals dominate such that inflation expectations respond negatively to contractionary policy, as would be expected given the transmission mechanism of monetary policy. However, we also find that inflation expectations respond positively to surprises in the Bank of England's macroeconomic projections, consistent with private agents putting more weight on the signal that they convey

[32] The Monetary Policy Committee has provided guidance on the setting of future monetary policy since August 7, 2013. Because this policy is supposed to affect the private agents' expected future policy path via a commitment device, it may affect private inflation expectations, and we need to control for this potential effect at the end of our sample.

[33] The wedge is computed as the difference between RPI and CPI inflation corrected for the uncertainty created by the announcement by the Office for National Statistics' Consumer Prices Advisory Committee (CPAC) of a potential revision in the RPI calculation methodology, between May 2012 and January 2013.

about future economic developments than the signal about the policy outlook. That provides evidence of the existence of a macro outlook signaling channel, in contrast to the theoretical predictions of full-information models. One interpretation of the empirical results could be that policy decisions and central bank inflation projections together enable private agents to differentiate the inflationary shock and the monetary shock, thus reducing the signaling effect of policy actions. The analysis of this hypothesis is left for future research. The results of this paper give policymakers some insights on how private agents interpret and respond to policy decisions and central bank information. The signals provided by central bank action and communication appear to be important for the management of private expectations.

References

Adam, K., and M. Padula. 2011. "Inflation Dynamics and Subjective Expectations in the United States." *Economic Inquiry* 49 (1): 13–25.

Andersen, T., T. Bollerslev, F. Diebold, and C. Vega. 2003. "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange." *American Economic Review* 93 (1): 38–62.

Andrade, P., G. Gaetano, E. Mengus, and B. Mojon. 2019. "Forward Guidance and Heterogeneous Beliefs." *American Economic Journal: Macroeconomics* 11 (3): 1–29.

Andrade, P., and H. Le Bihan. 2013. "Inattentive Professional Forecasters." *Journal of Monetary Economics* 60 (8): 967–82.

Angeletos, G.-M., C. Hellwig, and A. Pavan. 2006. "Signaling in a Global Game: Coordination and Policy Traps." *Journal of Political Economy* 114 (3): 452–84.

Baeriswyl, R., and C. Cornand. 2010. "The Signaling Role of Policy Actions." *Journal of Monetary Economics* 57 (6): 682–95.

Bean, C., and N. Jenkinson. 2001. "The Formulation of Monetary Policy at the Bank of England." *Quarterly Bulletin* (Bank of England) Winter (Q4): 434–41.

Beechey, M., B. Johannsen, and A. Levin. 2011. "Are Long-Run Inflation Expectations Anchored More Firmly in the Euro Area than in the United States?" *American Economic Journal: Macroeconomics* 3 (2): 104–29.

Bernanke, B., and A. Blinder. 1992. "The Federal Funds Rate and the Channels of Monetary Transmission." *American Economic Review* 82 (4): 901–21.

Bernanke, B., 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., and I. Mihov. 1998. "Measuring Monetary Policy." *Quarterly Journal of Economics* 113 (3): 869–902.

Blanchard, O., J.-P. L'Huillier, and G. Lorenzoni. 2013. "News, Noise, and Fluctuations: An Empirical Exploration." *American Economic Review* 103 (7): 3045–70.

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

Branch, W. 2004. "The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations." *Economic Journal* 114 (497): 592–621.

———. 2007. "Sticky Information and Model Uncertainty in Survey Data on Inflation Expectations." *Journal of Economic Dynamics and Control* 31 (1): 245–76.

Bullard, J., and K. Mitra. 2002. "Learning about Monetary Policy Rules." *Journal of Monetary Economics* 49 (6): 1105–29.

Campbell, J., C. Evans, J. Fisher, and A. Justiniano. 2012. "Macroeconomic Effects of Federal Reserve Forward Guidance." *Brookings Papers on Economic Activity* (Spring): 1–80.

Campbell, J., J. Fisher, A. Justiniano, and L. Melosi. 2017. "Forward Guidance and Macroeconomic Outcomes Since the Financial Crisis." *NBER Macroeconomics Annual 2016*, Vol. 31, ed. M. Eichenbaum and J. A. Parker, 283–357. University of Chicago Press.

Capistran, C., and A. Timmermann. 2009. "Disagreement and Biases in Inflation Expectations." *Journal of Money, Credit and Banking* 41 (2–3): 365–96.

Carroll, C. 2003. "Macroeconomic Expectations of Households and Professional Forecasters." *Quarterly Journal of Economics* 118 (1): 269–98.

Castelnuovo, E., and P. Surico. 2010. "Monetary Policy, Inflation Expectations and the Price Puzzle." *Economic Journal* 120 (549): 1262–83.

Cesa-Bianchi, A., G. Thwaites, and A. Vicondoa. 2020. "Monetary Policy Transmission in the United Kingdom: A High Frequency Identification Approach." *European Economic Review* 123 (April): 103–375.

Cieslak, A., and A. Schrimpf. 2019. "Non-Monetary News in Central Bank Communication." *Journal of International Economics* 118 (May): 293–315.

Cloyne, J., and P. Huertgen. 2016. "The Macroeconomic Effects of Monetary Policy: A New Measure for the United Kingdom." *American Economic Journal: Macroeconomics* 8 (4): 75–102.

Cochrane, J., and M. Piazzesi. 2002. "The Fed and Interest Rates—High-Frequency Identification." *American Economic Review* 92 (2): 90–95.

Coibion, O., and Y. Gorodnichenko. 2012. "What Can Survey Forecasts Tell Us about Informational Rigidities?" *Journal of Political Economy* 120 (1): 116–59.

———. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review* 105 (8): 2644–78.

Cook, T., and T. Hahn. 1989. "The Effect of Changes in the Federal Funds Rate Target on Market Interest Rates in the 1970s." *Journal of Monetary Economics* 24 (3): 331–51.

Crowe, C. 2010. "Testing the Transparency Benefits of Inflation Targeting: Evidence from Private Sector Forecasts." *Journal of Monetary Economics* 57 (2): 226–32.

D'Agostino, A., and K. Whelan. 2008. "Federal Reserve Information during the Great Moderation." *Journal of the European Economic Association* 6 (2–3): 609–20.

Dräger, L., M. Lamla, and D. Pfajfar. 2016. "Are Survey Expectations Theory-Consistent? The Role of Central Bank Communication and News." *European Economic Review* 85 (June): 84–111.

Ehrmann, M. 2015. "Targeting Inflation from Below: How Do Inflation Expectations Behave?" *International Journal of Central Banking* 11 (Supplement 1, September): 213–49.

Ehrmann, M., S. Eijffinger, and M. Fratzscher. 2012. "The Role of Central Bank Transparency for Guiding Private Sector Forecasts." *Scandinavian Journal of Economics* 114 (3): 1018–52.

Ehrmann, M., and M. Fratzscher. 2007. "Communication by Central Bank Committee Members: Different Strategies, Same Effectiveness?" *Journal of Money, Credit and Banking* 39 (2–3): 509–41.

Ellingsen, T., and U. Söderström. 2001. "Monetary Policy and Market Interest Rates." *American Economic Review* 91 (5): 1594–1607.

Evans, G., and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Princeton University Press.

Fatum, R., and M. Hutchison. 1999. "Is Intervention a Signal of Future Monetary Policy? Evidence from the Federal Funds Futures Market." *Journal of Money, Credit and Banking* 31 (1): 54–69.

Faust, J., E. Swanson, and J. Wright. 2004. "Do Federal Reserve Policy Surprises Reveal Superior Information about the Economy?" *Contributions to Macroeconomics* 4 (1).

Fendel, R., E. Lis, and J.-C. Rülke. 2011. "Do Professional Forecasters Believe in the Phillips Curve? Evidence from the G7 Countries." *Journal of Forecasting* 30 (2): 268–87.

Forbes, K., I. Hjortsoe, and T. Nenova. 2018. "The Shocks Matter: Improving Our Estimates of Exchange Rate Pass-Through." *Journal of International Economics* 114 (September): 255–75.

Fujiwara, I. 2005. "Is the Central Bank's Publication of Economic Forecasts Influential?" *Economics Letters* 89 (3): 255–61.

Gamber, E., and J. Smith. 2009. "Are the Fed's Inflation Forecasts Still Superior to the Private Sector's?" *Journal of Macroeconomics* 31 (2): 240–51.

Gürkaynak, R., A. Levin, and E. Swanson. 2010. "Does Inflation Targeting Anchor Long-Run Inflation Expectations? Evidence from the US, UK, and Sweden." *Journal of the European Economic Association* 8 (6): 1208–42.

Gürkaynak, R., B. Sack, and E. Swanson. 2005a. "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements." *International Journal of Central Banking* 1 (1, May): 55–93.

———. 2005b. “The Sensitivity of Long-term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models.” *American Economic Review* 95 (1): 425–36.

Gürkaynak, R., B. Sack, and J. Wright. 2010. “The TIPS Yield Curve and Inflation Compensation.” *American Economic Journal: Macroeconomics* 2 (1): 70–92.

Hansen, S., and M. McMahon. 2016. “Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication.” *Journal of International Economics* 99 (Supplement 1): S114–S133.

Hanson, S., and J. Stein. 2015. “Monetary Policy and Long-Term Real Rates.” *Journal of Financial Economics* 115 (3): 429–48.

Hayo, B., and M. Neuenkirch. 2010. “Do Federal Reserve Communications Help Predict Federal Funds Target Rate Decisions?” *Journal of Macroeconomics* 32 (4): 1014–24.

Hubert, P. 2014. “FOMC Forecasts as a Focal Point for Private Expectations.” *Journal of Money, Credit and Banking* 46 (7): 1381–1420.

———. 2015a. “Do Central Bank Forecasts Influence Private Agents? Forecasting Performance versus Signals.” *Journal of Money, Credit and Banking* 47 (4): 771–89.

———. 2015b. “Revisiting the Greenbook’s Relative Forecasting Performance.” *Revue de l’OFCE* 137: 151–79.

———. 2017. “Qualitative and Quantitative Central Bank Communication and Inflation Expectations.” *B.E. Journal of Macroeconomics* 7 (1): 1–41.

Jansen, D.-J., and J. De Haan. 2009. “Has ECB Communication Been Helpful in Predicting Its Interest Rate Decisions?” *Applied Economics* 41 (16): 1995–2003.

Jarocinski, M., and P. Karadi. 2020. “Deconstructing Monetary Policy Surprises: The Role of Information Shocks.” *American Economic Journal: Macroeconomics* 12 (2): 1–43.

Joyce, M., M. Tong, and R. Woods. 2011. “The United Kingdom’s Quantitative Easing Policy: Design, Operation and Impact.” *Quarterly Bulletin* (Bank of England) 51 (Q3): 200–212.

King, R., Y. Lu, and E. Pasten. 2008. “Managing Expectations.” *Journal of Money, Credit and Banking* 40 (8): 1625–66.

Kohlhas, A. 2014. “Learning-by-Sharing: Monetary Policy and the Information Content of Public Signals.” Manuscript.

Kuttner, K. 2001. "Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market." *Journal of Monetary Economics* 47 (3): 523–44.

Lyziak, T., and M. Paloviita. 2018. "On the Formation of Inflation Expectations in Turbulent Times: The Case of the Euro Area." *Economic Modelling* 72 (June): 132–39.

Mackowiak, B., and M. Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review* 99 (3): 769–803.

Mankiw, G., and R. Reis. 2002. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* 117 (4): 1295–1328.

Melosi, L. 2017. "Signalling Effects of Monetary Policy." *Review of Economic Studies* 84 (2): 853–84.

Miranda-Agrippino, S., and G. Ricco. 2020. "The Transmission of Monetary Policy Shocks." Forthcoming in *American Economic Journal: Macroeconomics.*

Moscarini, G. 2004. "Limited Information Capacity as a Source of Inertia." *Journal of Economic Dynamics and Control* 28 (10): 2003–35.

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

Nunes, R. 2010. "Inflation Dynamics: The Role of Expectations." *Journal of Money, Credit and Banking* 42 (6): 1161–72.

Orphanides, A., and J. Williams. 2005. "The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations." *Journal of Economic Dynamics and Control* 29 (11): 1927–50.

Paloviita, M., and M. Viren. 2013. "Are Individual Survey Expectations Internally Consistent?" Working Paper No. 140, National Bank of Poland.

Pedersen, M. 2015. "What Affects the Predictions of Private Forecasters? The Role of Central Bank Forecasts in Chile." *International Journal of Forecasting* 31 (4): 1043–55.

Reis, R. 2013. "Central Bank Design." *Journal of Economic Perspectives* 27 (4): 17–44.

Rigobon, R., and B. Sack. 2004. "The Impact of Monetary Policy on Asset Prices." *Journal of Monetary Economics* 51 (8): 1553–75.

Romer, C., and D. Romer. 2000. "Federal Reserve Information and the Behavior of Interest Rates." *American Economic Review* 90 (3): 429–57.

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

Rosa, C., and G. Verga. 2007. "On the Consistency and Effectiveness of Central Bank Communication: Evidence from the ECB." *European Journal of Political Economy* 23 (1): 146–75.

Rossi, B., and T. Sekhposyan. 2016. "Forecast Rationality Tests in the Presence of Instabilities, with Applications to Federal Reserve and Survey Forecasts." *Journal of Applied Econometrics* 31 (3): 507–32.

Scotti, C. 2016. "Surprise and Uncertainty Indexes: Real-Time Aggregation of Real-Activity Macro-Surprises." *Journal of Monetary Economics* 82 (September): 1–19.

Sims, C. 1992. "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy." *European Economic Review* 36 (5): 975–1000.

———. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50 (3): 665–90.

Soderlind, P. 2011. "Inflation Risk Premia and Survey Evidence on Macroeconomic Uncertainty." *International Journal of Central Banking* 7 (2, June): 113–33.

Sturm, J.-E., and J. De Haan. 2011. "Does Central Bank Communication Really Lead to Better Forecasts of Policy Decisions? New Evidence Based on a Taylor Rule Model for the ECB." *Review of World Economics* 147 (1): 41–58.

Swanson, E. 2006. "Have Increases in Federal Reserve Transparency Improved Private Sector Interest Rate Forecasts?" *Journal of Money, Credit and Banking* 38 (3): 792–819.

Tang, J. 2015. "Uncertainty and the Signaling Channel of Monetary Policy." Working Paper No. 15-8, Federal Reserve Bank of Boston.

Walsh, C. 2007. "Optimal Economic Transparency." *International Journal of Central Banking* 3 (1, March): 5–36.

Warsh, K. 2014. "Transparency and the Bank of England's Monetary Policy Committee." Technical Report, Bank of England (December).

Woodford, M. 2001. "Imperfect Common Knowledge and the Effects of Monetary Policy." NBER Working Paper No. 8673.

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

Optimal Monetary Policy in Small Open Economies: Producer-Currency Pricing*

Mikhail Dmitriev[a] and Jonathan Hoddenbagh[b]
[a]Florida State University
[b]Johns Hopkins University

We establish the share of exports in production as a sufficient statistic for optimal noncooperative monetary policy. Under financial autarky, markups positively co-move with the export share. For complete markets, markups should be procyclical if the export share is procyclical. When central banks cooperate, markups are constant under complete markets, and countercyclical under financial autarky.

JEL Codes: E50, F41, F42.

1. Introduction

Although price stability is widely viewed as a benchmark monetary policy for central banks, various ingredients in the open economy drive optimal policy away from replicating the flexible-price allocation. Cooperation in the absence of commitment (Rogoff 1985), imperfect risk sharing (Corsetti, Dedola, and Leduc 2010), incomplete exchange rate pass-through (Devereux and Engel 2003), noncooperative policy (De Paoli 2009a, 2009b), and trade elasticities (Benigno and Benigno 2003) generate deviations from price stability. Because there are so many additional ingredients in the open economy relative to the closed economy, it is very difficult to suggest a one-size-fits-all optimal monetary policy like price stability.

*This is a substantially revised and updated edition of the October 2014 and June 2013 version of the paper. We thank Fabio Ghironi, Susanto Basu, Pierpaolo Benigno, Giancarlo Corsetti, Eyal Dvir, Peter Ireland, and two anonymous referees for very helpful comments, as well as seminar participants at Boston College, the BC/BU Green Line Macro Meeting, the Eastern Economic Association Meeting, and the Canadian Economic Association Meeting. Any errors are our own. Author e-mails: mdmitriev@fsu.edu and jon.hoddenbagh@gmail.com.

The precise optimal policy is sensitive to a variety of assumptions and specific parameter settings.

Our goal is to provide a coherent, tractable framework to examine optimal monetary policy in small open economies under producer-currency pricing. We establish a set of simple rules to guide central banks in four unique cases where monetary policy is either cooperative or noncooperative and markets are complete or cross-border trade in financial assets is prohibited. These four cases nest most of the key distortions in the open economy: nominal rigidities, terms-of-trade externalities, and incomplete cross-country risk sharing.

For optimal noncooperative policy, we find that central banks should generate markups that follow the share of exported goods in total production, unless these markup movements cause excess consumption volatility. For example, if the share of exported goods is procyclical, then monetary policy should generate procyclical markups. For optimal cooperative policy, markups should be constant when markets are complete, and countercyclical under financial autarky.

The paper makes several contributions to the literature. First, we provide a unified framework for studying cooperative and noncooperative optimal monetary policy under both complete markets and financial autarky. Second, our solution is analytical and covers the full range of parameter values instead of focusing on a particular calibration. Third, we are the first to study cooperative policy for small open economies under producer-currency pricing (PCP) and show how it differs from the noncooperative case. Fourth, in our study of the optimal policy, we do not restrict import and export trade elasticities to be equal to each other. Finally, we establish the export share of gross domestic product (GDP) as a sufficient statistic for noncooperative policy.

We consider a continuum of small open economies that are hit by asymmetric productivity shocks, following Galí and Monacelli (2005). We deviate from their paper in three ways. First, we do not restrict our analysis to the widely used Cole-Obstfeld (1991) specification where the coefficient of relative risk aversion and trade elasticities are set to unitary values but instead analyze the most general case analytically. Second, we extend their analysis to both cooperative and noncooperative policies under financial autarky. Finally, we utilize one-period-in-advance price rigidities used by Obstfeld and

Rogoff (2000, 2002), Corsetti and Pesenti (2001, 2005), Faia and Monacelli (2008), Dmitriev and Hoddenbagh (2019), and Egorov and Mukhin (2019) instead of a more traditional Calvo setup.

We frame monetary policy in terms of deviations from the flexible-price allocation using markups. For example, when the policymaker intends to decrease markups, she lowers the interest rate and depreciates the currency. Producer prices remain stable in the home currency and fall when expressed in foreign currencies. Thus, export volume increases. At the same time, import prices remain constant in foreign currencies and rise in the home currency. Therefore, the terms of trade depreciate. Also, locally produced goods become more competitive at home and crowd out imports in terms of volumes. Consumer prices, composed of higher import prices and stable local product prices, increase. Finally, output and employment, driven by higher demand for exports and domestic import substitution, tend to go up, raising wages and reducing markups. Thus, negative deviations of markups from the steady state are associated with expansionary monetary policy, which generates positive output gaps, currency depreciation, and price and wage inflation.

We begin our analysis by considering cooperative policy and complete markets. In this case, nominal rigidity is the only distortion present, and optimal monetary policy replicates the flexible-price allocation through a policy of price stability. While we are the first to consider cooperative policy under complete markets for small open economies, our contribution for this specification is mainly technical. For example, Benigno and Benigno (2006) and Corsetti, Dedola, and Leduc (2010) have established that replicating the flexible-price allocation is optimal for cooperative policymakers under complete markets for two large open economies.

We next consider optimal cooperative policy under financial autarky. In this case, there are two distortions: nominal rigidities and incomplete risk sharing across countries. For empirically relevant parameter settings, we show that optimal cooperative monetary policy should generate countercyclical markups. These markup movements are designed to manipulate the terms of trade and redistribute resources from countries with positive supply shocks to countries hit by adverse shocks via terms-of-trade depreciation for the former and appreciation for the latter. As trade elasticities rise and monopoly power at the export level deteriorates, central banks lose their ability to influence the terms of trade, such that they focus more on replicating flexible prices and less on terms-of-trade adjustments.

The closest relevant study by Corsetti, Dedola, and Leduc (2010) is for large open economies. They set import and export elasticities to be equal. While their focus is when optimal policy replicates the flexible-price allocation or the first-best allocation, we study the full markup dynamics.

Next, we analyze optimal policy under financial autarky without cooperation. There are three distortions that drive the equilibrium away from the efficient allocation: nominal rigidities, terms-of-trade externalities, and market incompleteness. We give an explicit analytic expression for the optimal markup, and then establish the share of exports in production as a sufficient statistic for the optimal monetary policy. Optimal markups positively co-move with the export share. If the export share is constant, replication of the flexible-price allocation is optimal. Indeed, under noncooperative policy, risk sharing across countries becomes irrelevant for the policymaker. The policymaker desires to sell exports abroad with a positive markup and have no markup for products produced and consumed at home. Moreover, when prices are set one period in advance, policymakers cannot influence markups systematically. As a result, central banks tend to increase (decrease) markups when the export share goes up (down).

De Paoli (2009b) also analyzes noncooperative policy under financial autarky for the limiting case of two large open economies. We differ from her analysis in several ways. First, her study has a more quantitative focus so that she fixes most of the parameters to particular values and uses Calvo pricing. Instead, we consider markup movements for the broadest possible range of parameter values. Second, we provide a sufficient statistic for optimal monetary policy: the policymaker only needs the dynamics of the export share regardless of the underlying parameters. Third, we allow export and import elasticities to differ from each other.

Finally, we consider the case of complete markets and noncooperative policy. In this setting the optimal markup is procyclical whenever the export share is procyclical. A procyclical markup enables the policymaker to extract higher monopolistic rents from foreigners through terms-of-trade appreciation which stabilizes domestic consumption. On the other hand, when the export share is countercyclical, countercyclical markups generate stronger terms-of-trade externality rents and destabilize consumption. When the costs from destabilizing consumption exceed the benefits from the terms-of-trade externality, the optimal markups might be procyclical despite the countercyclical export share.

Noncooperative policy under complete markets and PCP has been studied by Faia and Monacelli (2008) and De Paoli (2009a, 2009b) for small open economies as the limiting case of two large economies. As mentioned before, we differ from these studies by differentiating between trade elasticities and by deriving analytical expressions for markups and other variables, instead of focusing on a particular calibration with Calvo pricing. We also consider a case where trade elasticities equal to each other.

Using the export share as a sufficient statistic helps to explain why under the Cole-Obstfeld (1991) specification, where trade elasticities and risk aversion are set to one, the flexible-price allocation is optimal under all four cases considered here. Under Cole-Obstfeld the export share is constant and terms-of-trade movements provide complete risk sharing. As a result, policymakers have no incentive to stabilize consumption or extract monopolistic rents from foreigners, as the export share is constant in all cases. Also, our principle explains why under noncooperative policy the flexible-price allocation is optimal for fully open economies as the export share remains equal to one over the cycle. The export share becomes less relevant for noncooperative policy only if the central bank loses its ability to influence the terms of trade. Then the optimal policy is to replicate the flexible-price allocation.

There are several reasons why we use producer-currency pricing, despite some recent empirical evidence supporting dominant-currency pricing, or DCP (Goldberg and Tille 2008; Gopinath, Itskhoki, and Rigobon 2010; Gopinath et al. 2016). First, both PCP and DCP imply equal sensitivity of import prices to the exchange rate. Second, the evidence by Amiti, Itskhoki, and Konings (2014) on export prices supports at least 50 percent exchange rate pass-through for large import-intensive exporters and full pass-through for smaller exporters, while DCP assumes no pass-through for export prices. Third, under PCP, high export elasticities allow terms of trade to be stable and independent from monetary policy, similar to

DCP. Finally, PCP, formally developed by Fleming (1962), Mundell (1963), Obstfeld and Rogoff (1995), and others, serves as the benchmark case for optimal monetary policy analysis in the open economy since other setups add extra distortions in addition to the ones present in PCP.

We differ from many papers in the field by having prices set one period in advance instead of Calvo pricing. This price setting allows us to arrive at the optimal conditions in a fully nonlinear manner. We linearize the equilibrium afterward, which makes our results robust to the Kim and Kim (2003) critique of linear-quadratic approximation.[1] Also, the Calvo approach to nominal rigidities introduces price dispersion in addition to standard output gap costs. Although price dispersion increases the complexity of the model, it is proportional to the output gap, which makes economic intuition for optimal policy similar between Calvo pricing and one-period-in-advance pricing. As our contribution is more analytical than quantitative, we use one-period-in-advance price setting for the sake of tractability.

We also differ from the optimal policy literature by studying a continuum of small open economies instead of considering a limit of two economies. The continuum allows us to differentiate import and export elasticities, which are equal in the case of two countries. In the absence of this differentiation, the analysis faces empirical and theoretical difficulties. For example, under inelastic export demand, an increase in the quantity of exports leads to lower export revenues, which makes positive supply shocks potentially welfare reversing. Second, under inelastic export demand with constant elasticity, an infinite export tax generates limitless export revenues and consumption.[2] To avoid these problems, we set the export elasticity to be greater than one, and allow the import elasticity to vary between zero and infinity. Forcing import and export elasticities to be equal allows us to nest the limiting case of two open economies. Also, a continuum of small open economies under cooperation nests two large open economies as a special case. For instance, if we divide the

[1]Note that Benigno and Woodford (2012) establish conditions under which linear-quadratic analysis is robust to the Kim and Kim (2003) critique.

[2]Infinite export tariffs are never optimal for two large open economies, where they generate a negative wealth effect through impoverishing the trading partner and provoking trade wars. However, potential welfare reversals under inelastic export demand are still present even for two large open economies.

countries in a continuum into two groups and allow them to have two realizations of the technology shocks, the framework is equivalent to two large open economies, as small open economies differ only by the realization of the technology shock. While we do not introduce symmetric shocks, the division of the asymmetric shocks into two groups of positive and negative values allows for a close approximation of two large open economies.[3]

Our research is also closely related to the old debate on the value of trade elasticities in the data. While the trade literature finds larger elasticities, with values ranging between four and five (Anderson and van Wincoop 2004; Simonovska and Waugh 2014a, 2014b; Imbs and Mejean 2015), the international macro and finance literature (a non-exhaustive list starting from Backus, Kehoe, and Kydland 1994; Stockman and Tesar 1995; and many others) often assumes these elasticities are smaller, with values ranging between 0.8 and 1.5. This debate has a strong effect on optimal monetary policy where trade elasticities play a central role. We show that this debate is not relevant for the policymaker since the effect of trade elasticities on monetary policy is expressed through the dynamics of the directly observed export share.

2. The Model

We consider a continuum of small open economies represented by the unit interval, as popularized in the literature by Galí and Monacelli (2005, 2008).[4] Each economy consists of a representative household and a representative firm. All countries are identical ex ante: they have the same preferences, technology, and price setting. Ex post, economies differ depending on the realization of their technology shock. Households are immobile across countries, while goods can move freely across borders. Each economy produces one final good, over which it exercises a degree of monopoly power. In particular, countries are able to manipulate their terms of trade even though

[3]Noncooperative policy for two large open economies differs from the noncooperative policy under a continuum due to the presence of strategic interactions.

[4]A similar version of this model appears in Dmitriev and Hoddenbagh (2019), where we employ wage rigidity instead of price rigidity and study the optimal design of a fiscal union within a currency union.

they are measure zero, similar to an individual producer in a model of monopolistic competition. However, because countries are small, they have no impact on world income or the world interest rate.

We use a one-period-in-advance price setting to introduce nominal rigidities. Monopolistic firms set the next period's nominal prices in terms of the domestic currency, before the next period's production and consumption decisions. These firms charge a constant markup in the flexible-price equilibrium, utilizing their monopoly power at the firm level. Given this preset price, firms supply as much output as demanded by households.

We lay out a general framework below and then focus on four particular cases: cooperative policy under complete markets and financial autarky, and noncooperative policy under complete markets and financial autarky.

Households. In each economy $i \in [0, 1]$, there is a representative household with lifetime expected utility

$$\mathbb{E}_t \left\{ \sum_{k=0}^{\infty} \beta^k \left(\frac{C_{it+k}^{1-\sigma}}{1-\sigma} - \frac{N_{it+k}^{1+\varphi}}{1+\varphi} \right) \right\}, \tag{1}$$

where $\beta < 1$ is the household discount factor, C is the consumption basket, and N is household labor effort. Households face a general budget constraint that nests both complete markets and financial autarky; we will discuss the differences between the two in subsequent sections. For now, it is sufficient to simply write out the most general form of the budget constraint:

$$C_{it} = (1+\tau_i)\left(\frac{W_{it}}{P_{it}}\right) N_{it} + \frac{P_{F,it}}{P_{it}} \mathcal{D}_{it} + \Pi_{it} - \mathcal{T}_{it} + \frac{B_{it}}{P_{it}} - \frac{B_{it+1}}{R_{it} P_{it}}. \tag{2}$$

The distortionary subsidy on household labor income in country i is denoted by τ_i, while $\mathcal{T}_{it}$ is a lump-sum tax collected from the households to finance this subsidy. Overall, the government budget is balanced at every period. These subsidies and taxes are designed to enforce the efficient steady-state allocation. Π_{it} denotes profits from the monopolistic firms which are distributed lump sum to households. Without loss of generality, we assume that equities in the model are not traded. The consumer price index corresponds to P_{it}, while the nominal wage is W_{it}. The price index $P_{F,it}$ reflects the price of the basket composed from the imported goods expressed in units of currency i. $\mathcal{D}_{it}$ denotes net state-contingent portfolio payments expressed in the units of the imported goods, which are available to households under complete markets. This portfolio consists of state-contingent bonds, available for every state of the world. To simplify the exposition, we allow only domestic households to hold noncontingent bonds B_{it}. This is not a limitation, as foreigners have access to the portfolio of state-contingent securities. In equilibrium, noncontingent bonds are relevant only as a source of information to pin down interest rate dynamics. When international asset markets are complete, households perform all cross-border trades in contingent claims in period 0, insuring against all possible states in all future periods. Under financial autarky, households have access only to noncontingent bonds so that $\mathcal{D}_{it} = 0$.

Consumption and Price Indexes. The consumption basket for a representative small open economy i consists of home goods, $C_{H,it}$, and foreign goods, denoted by $C_{F,it}$. It is defined as follows:

$$C_{it} = \left[(1-\alpha)^{\frac{1}{\eta}}(C_{H,it})^{\frac{\eta-1}{\eta}} + \alpha^{\frac{1}{\eta}}(C_{F,it})^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}, \tag{3}$$

where the import basket $C_{F,it}$ is defined as

$$C_{F,it} = \left(\int_0^1 (C_{F,ijt})^{\frac{\gamma-1}{\gamma}} dj\right)^{\frac{\gamma}{\gamma-1}}. \tag{4}$$

The variable $C_{F,ijt}$ corresponds to consumption in country i of the variety produced in country j. The value of η reflects the elasticity of substitution between domestically produced goods and imported varieties in the aggregate consumption basket. The parameter γ corresponds to the elasticity of substitution between exported domestic and foreign exported varieties. The degree of consumption home bias is represented by the value $1-\alpha$. Therefore, in a fully closed economy $\alpha = 0$, and in the fully open economy $\alpha = 1$.

The consumer price index P_{it} is an aggregator of the domestic variety price $P_{H,it}$ and import price index $P_{F,it}$:

$$P_{it} = \left[(1-\alpha)P_{H,it}^{1-\eta} + \alpha P_{F,it}^{1-\eta}\right]^{\frac{1}{1-\eta}}. \tag{5}$$

Here the import price index $P_{F,it}$ is a constant elasticity of substitution aggregator that takes the following form:

$$P_{F,it} = \left(\int_0^1 (\mathcal{E}_{ijt} P_{H,jt})^{1-\gamma} dj \right)^{\frac{1}{1-\gamma}}, \tag{6}$$

where the variety produced in country j is sold in country i for a price $\mathcal{E}_{ijt} P_{H,jt}$. We define the nominal bilateral exchange rate $\mathcal{E}_{ijt}$ as units of currency i per one unit of currency j.

Household expenditure minimization yields the demand for the home variety $C_{H,it}$, the demand for imported varieties $C_{F,it}$, and the relative demand for the variety produced in country j and consumed in country i:

$$C_{H,it} = (1-\alpha) \left(\frac{P_{H,it}}{P_{it}} \right)^{-\eta} C_{it}, \tag{7}$$

$$C_{F,it} = \alpha \left(\frac{P_{F,it}}{P_{it}} \right)^{-\eta} C_{it}, \tag{8}$$

$$C_{F,ijt} = \left(\frac{P_{H,jt}}{P_{F,jt}} \right)^{-\gamma} C_{F,it}. \tag{9}$$

We assume that producer-currency pricing holds so that the law of one price applies. Put differently, the price of the same good is equal across countries when converted into a common currency. In this case, the good produced in country i has a price in country j equal to $\mathcal{E}_{jit} P_{H,it}$. Although the law of one price holds for individual varieties, purchasing power parity does not hold because of home bias in consumption. Nevertheless, import price indexes are identical across countries when converted to the same currency, such that $P_{F,it} = \mathcal{E}_{ijt} P_{F,jt}$. The terms of trade for country j, defined as the home-currency price of exports over the home-currency price of imports, is denoted $\tilde{P}_{H,jt} = \frac{P_{H,jt}}{P_{F,jt}}$. We define the aggregate consumer price index normalized by the import price index as $\tilde{P}_{it} = \frac{P_{it}}{P_{F,it}}$. In our model, log-deviations of $\tilde{P}_{it}$ from the steady state correspond to real exchange rate movements to a first-order approximation. Using the normalized price levels $\tilde{P}_{H,it}$ and $\tilde{P}_{it}$, we modify the demand expressions (10), (11), and (12) into

$$C_{H,it} = (1-\alpha)\tilde{P}_{H,it}^{-\eta}\tilde{P}_{it}^{\eta}C_{it}, \tag{10}$$

$$C_{F,it} = \alpha C_{it}\tilde{P}_{it}^{\eta}, \tag{11}$$

$$C_{F,ijt} = \tilde{P}_{H,it}^{-\gamma}C_{F,it}. \tag{12}$$

Goods market clearing requires that the supply of the domestic variety produced in country i equals demand from home consumers $C_{H,it}$ and foreign consumers $\int_0^1 C_{F,jit}dj$:

$$Y_{it} = C_{H,it} + \int_0^1 C_{F,jit}dj. \tag{13}$$

Substituting equations (10) and (12) into the goods market clearing condition (13) yields global demand for country i's unique variety:

$$Y_{it} = (1-\alpha)\left(\frac{P_{H,it}}{P_{it}}\right)^{-\eta} C_{it} + \left(\frac{P_{H,it}}{P_{F,it}}\right)^{-\gamma}\int_0^1 C_{F,jt}dj. \tag{14}$$

Given the symmetric structure of the model as well as the independence of idiosyncratic shocks across countries, the integral on the import basket in (14) is equivalent to the unconditional expectation, which corresponds to the ergodic mean of the import basket. More formally,

$$\int_0^1 C_{F,jt}dj = \mathbb{E}\{C_{F,it}\} = \alpha\mathbb{E}\left\{\left(\frac{P_{F,it}}{P_{it}}\right)^{-\eta} C_{it}\right\} = \alpha\mathbb{E}\{\tilde{P}_{it}^{\eta}C_{it}\}. \tag{15}$$

Noncooperative policymakers take global import $\int_0^1 C_{F,jt}dj$ as given. Substituting (15) into (14) yields the following goods market clearing condition:

$$Y_{it} = (1-\alpha)\tilde{P}_{H,it}^{-\eta}\tilde{P}_{it}^{\eta}C_{it} + \alpha\tilde{P}_{H,it}^{-\gamma}\mathbb{E}\{\tilde{P}_{it}^{\eta}C_{it}\}. \tag{16}$$

We define the share of the domestically produced variety in economy i that is exported goods as

$$E_{s,it} = \frac{\alpha\tilde{P}_{H,it}^{-\gamma}\mathbb{E}\{\tilde{P}_{it}^{\eta}C_{it}\}}{Y_{it}}. \tag{17}$$

Production. Each economy i consists of a group of intermediate goods producers, $h \in [0,1]$, who exercise monopoly power over their unique variety, and a perfectly competitive final goods producer, who aggregates the intermediates in a constant elasticity of substitution fashion into a final good. For simplicity, we assume that intermediates are nontradable. Thus, each country bundles its intermediates into one final good, which is consumed both at home and abroad.[5]

Production of intermediates requires technology Z_{it}, which is common across firms within a country, and labor $N_{it}(h)$, which is unique to each firm. We assume that technology is independent across time and across countries (assumptions that can be easily relaxed) but is identical across firms within the same country. Given this, the production function of a representative intermediate goods firm h in country i is $y_{it}(h) = Z_{it}N_{it}(h)$, and aggregate output is described by

$$Y_{it} = Z_{it}N_{it}. \tag{18}$$

Because intermediate goods firms produce differentiated varieties, they exercise monopoly power and charge markups over their costs. A perfectly competitive final goods producer aggregates the intermediate input of each firm in the following way:

$$Y_{it} = \left[\int_0^1 Y_{it}(h)^{\frac{\varepsilon-1}{\varepsilon}} dh\right]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{19}$$

where ε is the elasticity of substitution between different intermediates. For country i, the price of the final good, $P_{H,it}$, is a

[5] We assume nontradable intermediates with a final tradable consumption good that aggregates those intermediates for simplicity. In Galí and Monacelli's (2005, 2008) setup, intermediate goods are tradable, such that every country's import consumption basket is made up of an infinite number of varieties imported from an infinite number of countries. This assumption requires integrating over two continuums. While it is straightforward for us to maintain their setup, we prefer the tractable alternative: a final goods producer bundles the domestically produced intermediates for export. In this way, each country produces only one unique variety, and we only need to integrate over one continuum. This assumption does not change the results in any way. In both cases the household consumption basket in each country is made up of imported goods from all i countries, which are themselves made up of intermediates produced domestically.

function of the nominal price for intermediate goods, $P_{H,it}(h)$: $P_{H,it} = \left[\int_0^1 P_{H,it}(h)^{1-\varepsilon} dh\right]^{\frac{1}{1-\varepsilon}}$. Cost minimization by the perfectly competitive final goods exporter leads to the following demand for intermediate variety h:

$$Y_{it}(h) = \left[\frac{P_{H,it}(h)}{P_{H,it}}\right]^{-\varepsilon} Y_{it}. \tag{20}$$

Intermediate goods firms choose the profit-maximizing price for their unique good one period in advance according to the following condition:

$$P_{H,it}(h) = \mu \frac{\mathbb{E}_{t-1}\left\{C_{it}^{-\sigma} Y_{it}(h) \frac{W_{it}}{Z_{it}P_{it}}\right\}}{\mathbb{E}_{t-1}\left\{\frac{C_{it}^{-\sigma} Y_{it}(h)}{P_{it}}\right\}}, \tag{21}$$

where $\mu = \frac{\varepsilon}{\varepsilon - 1}$ is the markup, which defines the degree of monopoly power at the firm level.

Households maximize utility (1) subject to their budget constraint (2). The first-order condition with respect to labor gives the following household labor supply:

$$\frac{W_{it}}{P_{it}} = \frac{1}{1+\tau} N_{it}^{\varphi} C_{it}^{\sigma}. \tag{22}$$

Because firms are identical at the national level, in equilibrium $P_{H,it}(h) = P_{H,it}$ and $Y_{it}(h) = Z_{it}N_{it}(h) = Z_{it}N_{it} = Y_{it}$. To eliminate steady-state markups at the firm level and achieve the first-best steady state, we choose distortionary labor subsidies such that $\frac{\mu}{1+\tau} = 1$. Using the optimal pricing equation (21), the labor supply condition (22), and the fact that prices are preset at time $t-1$, the optimal pricing condition under PCP is

$$1 = \frac{\mathbb{E}_{t-1}\left\{N_{it}^{1+\varphi}\right\}}{\mathbb{E}_{t-1}\left\{C_{it}^{-\sigma} \frac{N_{it}Z_{it}\tilde{P}_{H,it}}{\tilde{P}_{it}}\right\}}. \tag{23}$$

We evaluate monetary policy in relation to markup fluctuations. In our setup, flexible prices are equivalent to constant markups. As marginal costs for one unit of the final good are equal to $\frac{W_{it}}{Z_{it}}$, we define the markup as

$$\mu_{it} = \frac{P_{H,it}}{\frac{W_{it}}{Z_{it}}}. \tag{24}$$

Substituting the labor supply condition (22), the terms of trade $\tilde{P}_{H,it} = \frac{P_{H,it}}{P_{F,it}}$ into (24), we obtain

$$\mu_{it} = \frac{Z_{it}\frac{\tilde{P}_{H,it}}{\tilde{P}_{it}}}{N_{it}^{\varphi}C_{it}^{\sigma}}. \tag{25}$$

Note, that under flexible prices we set $\mu_{it} = 1$, which fully corresponds to the optimal pricing condition (23) once we take out the expectations. The intuitive interpretation of (23) is that firms set up prices equal to expected marginal costs. Optimal pricing condition can also be formulated in terms of markups. Plugging (25) into (23) gives

$$1 = \frac{\mathbb{E}_{t-1}\{N_{it}^{1+\varphi}\}}{\mathbb{E}_{t-1}\{N_{it}^{1+\varphi}\mu_{it}\}}. \tag{26}$$

Intuitively, equation (26) says that markups on average are equal to one. In other words, the central bank chooses cyclical markup dynamics around one to maximize household welfare subject to market clearing constraints.

2.1 Complete Markets

In complete markets, agents in each economy have access to a full set of domestic and foreign state-contingent assets to insure against country-specific consumption risk. Households in all countries maximize their lifetime utility (1) choosing consumption, labor, and a complete set of nominal state-contingent portfolio payments, subject to the budget constraint (2). Since countries are symmetric ex ante, complete markets imply the following risk-sharing condition:

$$\frac{C_{it}^{-\sigma}}{\tilde{P}_{it}} = \frac{C_{jt}^{-\sigma}}{\tilde{P}_{jt}} \quad \forall i, j, \tag{27}$$

which states that the marginal utility from consumption of imported varieties $C_{F,it}$, which is equal to the ratio of marginal utility of consumption and normalized aggregate price index, must be equal across all countries.

When international asset markets are complete, households perform all cross-border trades in contingent claims in period 0 before the realization of any shocks, insuring against all possible states in all future periods. To ensure that there are no Ponzi schemes in issuing state-contingent securities, we impose the intertemporal asset constraint that all transactions in period 0 before the realization of shocks must be balanced. Payment for claims issued in subsequent periods must equal payment for claims received. In online appendix A (available at http://www.ijcb.org), we show that the intertemporal asset constraint for complete markets is

$$\mathbb{E}\left\{\sum_{t=1}^{\infty}\beta^t C_{it}^{-\sigma}\frac{\mathcal{D}_{it}}{\tilde{P}_{it}}\right\} = 0, \tag{28}$$

which corresponds to equation (A.6). Intuitively, the intertemporal asset constraint stipulates that the present discounted value of future earnings should be equal to the present discounted value of future consumption flows. Total state-contingent portfolio payments across countries in every period must sum to zero such that in the absence of worldwide uncertainty

$$\mathbb{E}\{\tilde{P}_{it}C_{it}\} = \mathbb{E}\{\tilde{P}_{H,it}Y_{it}\}. \tag{29}$$

In online appendix A we combine the risk-sharing condition (27) and balanced portfolio flows among all countries in the absence of symmetric shocks (29) to yield the following expression for consumption in country i:

$$\mathbb{E}\left\{\tilde{P}_{it}^{\frac{\sigma-1}{\sigma}}\right\}\tilde{P}_{it}^{\frac{1}{\sigma}}C_{it} = \mathbb{E}\left\{\tilde{P}_{H,it}Y_{it}\right\}, \tag{30}$$

which corresponds to (A.9).

Our treatment of complete markets appears somewhat different from the rest of the literature. For example, Galí and Monacelli (2005) do not use the intertemporal asset constraint (28). As a result, they assume that the value $\frac{C_{it}^{-\sigma}}{\tilde{P}_{it}}$ is constant and independent from monetary policy, which is not correct. While the expression $\frac{C_{it}^{-\sigma}}{\tilde{P}_{it}}$ is indeed independent of the realization of the shocks, our derivations show that it is a composite of ergodic means of the endogenous variables, which are not only affected by the monetary policy but can potentially have a first-order impact on the optimal policy rule itself.

2.2 Financial Autarky

In financial autarky there is no trade in state-contingent financial assets, such that $\mathcal{D}_{it} = 0 \;\; \forall i, t$. The aggregate resource constraint under financial autarky specifies that the nominal value of output in the home country must equal the nominal value of consumption in the home country:

$$P_{it} C_{it} = P_{H,it} Y_{it}. \tag{31}$$

Normalizing this expression by $P_{F,it}$ gives

$$\tilde{P}_{it} C_{it} = \tilde{P}_{H,it} Y_{it}. \tag{32}$$

2.3 Technology Shocks

We assume that technology shocks are independent and identical across time and countries, and have a log-normal distribution such that

$$\log Z_{it} \sim \mathcal{N}(0, \sigma_z^2). \tag{33}$$

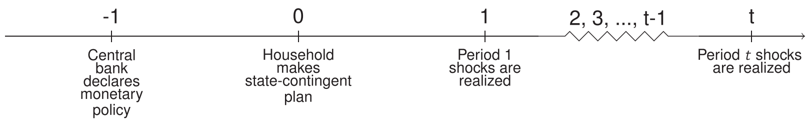
The model is formulated such that independence across time can be relaxed. It is also straightforward to relax the assumption of independence across countries by introducing a global aggregate component to technology. Our conclusions are robust to assumption of iid exogenous shock dynamics.

3. Optimal Monetary Policy

3.1 Setting Up the Optimization Problem

Without loss of generality, we assume a cashless limiting economy. Central banks use an interest rate rule to set monetary policy,

Figure 1. The Timeline of the Model



which affects the dynamics of the nominal exchange rate through uncovered interest rate parity. Under PCP, fluctuations in the nominal exchange rate pass through fully to import prices expressed in domestic currency $P_{F,it}$. As the price of the final domestic good is fixed one period in advance, nominal exchange rate fluctuations affect the terms of trade $\tilde{P}_{H,it}$. The terms of trade, in turn, affect the real exchange rate, consumption, and hours worked. Ultimately, a change in hours worked affects household disutility from labor, wages, and markups. We refer the reader to online appendix D, where the relationship between the interest rate and the terms of trade is formally established.

The timing of the model is described in figure 1. Before any shocks are realized, national central banks declare their policy for all states of the world. With this knowledge in hand, households lay out a state-contingent plan for consumption, labor hours, money, and asset holdings. After that, shocks hit the economy. Note that under financial autarky, no international asset trading occurs.

We summarize the optimization problem for the four cases we consider below. In each economy, the central bank maximizes the utility[6] of the representative household

$$\max_{\tilde{P}_{H,t}, P_t, C_t, N_t} \mathbb{E}\left\{\sum_{t=1}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\varphi}}{1+\varphi}\right]\right\}, \tag{34}$$

subject to the optimal pricing condition (35), goods market clearing (36), aggregate consumer price index (37), and asset market clearing (38). The first-order conditions with respect to the terms of trade

[6] Pricing one period in advance allows us to derive fully nonlinear first-order conditions without relying on linear-quadratic approximations, similar to Egorov and Mukhin (2019).

$\tilde{P}_{H,t}$[7], the normalized price index $\tilde{P}_t$, consumption C_t, and labor N_t are

$$\mathbb{E}_{t-1}\left\{C_t^{-\sigma}\frac{N_t Z_t \tilde{P}_{H,t}}{\tilde{P}_t}\right\} = \mathbb{E}_{t-1}\left\{N_t^{1+\varphi}\right\}, \tag{35}$$

$$Z_t N_t = (1-\alpha)\tilde{P}_{H,t}^{-\eta} C_t P_t^{\eta} + \alpha \tilde{P}_{H,t}^{-\gamma}(\mathbb{1}_{CP}\mathbb{E}[C_t P_t^{\eta}] + (1-\mathbb{1}_{CP})C_F), \tag{36}$$

$$\tilde{P}_t^{1-\eta} = (1-\alpha)\tilde{P}_{H,t}^{1-\eta} + \alpha, \tag{37}$$

$$C_t = \mathbb{1}_{CM}\frac{\mathbb{E}\{Z_t N_t \tilde{P}_{H,t}\}}{\tilde{P}_t^{\frac{1}{\sigma}}\mathbb{E}[\tilde{P}_t^{\frac{\sigma-1}{\sigma}}]} + (1-\mathbb{1}_{CM})\frac{Z_t N_t \tilde{P}_{H,t}}{\tilde{P}_t}. \tag{38}$$

Without loss of generality subscript i is omitted in expressions (34)–(38), as the economies effectively differ only by the realization of the technology shock. The indicator function $\mathbb{1}_{CP}$ is equal to one when the policy is cooperative, and zero otherwise. The indicator function $\mathbb{1}_{CM}$ is equal to one when financial markets are complete, and zero under financial autarky. When the policy is cooperative, the domestic central bank takes into account the effect its policy on other countries through the average consumption of foreign goods $\int_0^1 C_{F,it}di = \mathbb{E}C_{F,it} = \mathbb{E}[C_t\tilde{P}_t]$. On the other hand, when the policy is noncooperative, central banks take aggregate world consumption of foreign goods as given so that $\int_0^1 C_{F,it} = C_F$.

Stochastic processes for the terms of trade $\tilde{P}_{H,t}$ and the technology shock Z_t fully define the equilibrium path for the endogenous variables $\tilde{P}_t$, C_t, N_t using (35)–(38). Under the optimal monetary policy, the terms of trade $\tilde{P}_{H,t}$ reacts endogenously to the technology shock Z_t so that all variables $\tilde{P}_{H,t}$, $\tilde{P}_t$, C_t, N_t can be expressed as

[7] As mentioned at the beginning of this section, the central bank affects the terms-of-trade path only indirectly through its choice of the interest rate path, which in turn affects the nominal exchange rate through uncovered interest rate parity. Under producer-currency pricing, the price of home exports is set one period in advance in domestic currency, while import prices move with the exchange rate, changing the terms of trade. The equivalence between the allocation resulting from direct central bank control of the terms of trade and the allocation resulting from the central bank choosing an interest rate path is established in online appendix D.

functions of the technology shock. Before we analyze optimal monetary policy, let us consider a few stochastic processes for the terms of trade.

3.2 Stochastic Processes for the Terms of Trade

Traditionally, monetary policy in closed economies is expressed using nominal interest rate rules such as the Taylor rule. In our model, the interest rate affects the rest of the economy through the terms of trade. For analytical convenience, we consider shocks to the terms of trade instead of shocks to the interest rate rule. In particular, we employ a class of interest rate paths that generate the following stochastic process for the terms of trade:

$$\tilde{P}_{H,t} = f(Z_t). \tag{39}$$

We focus on interest rate rules that are not history dependent, where the terms of trade react only to the technology shock in the current period. The log-linear equivalent of the process described by (39) is

$$\hat{\tilde{P}}_{H,t} = \alpha_T \hat{Z}_t. \tag{40}$$

After we combine (39) with the constraints (35)–(38), we can express the dynamics of the log-linearized endogenous variables $(\hat{Y}_t, \hat{C}_t, \hat{N}_t, \hat{\tilde{P}}_t, \hat{\mu}_t)$ in terms of the technology shock. The search for the optimal monetary policy rule is equivalent to searching for the coefficient α_T that gives the highest welfare. To compare different policy rules and values for α_T, we need to know how terms-of-trade fluctuations affect other endogenous variables for a given technology shock. Lemma 1 below summarizes the impact of the terms of terms of trade on other key endogenous variables for a given technology shock.

LEMMA 1. *For an equilibrium that is characterized by (35)–(39) and for a given realization of the technology shock, a monetary policy that generates lower (higher) terms of trade also leads to lower (higher) markups, and higher (lower) consumption, employment, and output. This result holds under financial autarky and complete markets, for cooperative and noncooperative policies.*

Proof. See online appendix C.1. ∎

Lemma 1 describes the effect of the terms-of-trade depreciation on the economy for a given technology shock. The transmission mechanism from expansionary monetary policy through terms-of-trade depreciation into markups, consumption, hours worked, and output is intuitive. Lower terms of trade lead to an increase in consumption, employment, and output through higher exports and import substitution, as currency depreciation makes domestic goods more competitive at home and abroad. Higher demand for labor drives up wages, and in the presence of price rigidities reduces markups.

It is natural to consider the stochastic process for the terms of trade that can replicate the flexible-price allocation.

LEMMA 2. *Under the flexible-price allocation, a positive technology shock increases consumption and output and leads to a decline in the terms of trade and a real exchange rate depreciation. This result holds under both financial autarky and complete markets.*

Proof. See online appendix C.2. ∎

Lemma 2 states that under flexible prices, a positive technology shock increases consumption and output, similar to a closed economy. An increase in productivity also leads to a decrease in the terms of trade as abundant home-produced goods become cheaper relative to foreign goods.

The optimal monetary policy under some circumstances differs from replication of the flexible-price allocation. We use markup dynamics to characterize monetary policy, while in the literature monetary policy is often characterized by the importance of flexible prices or exchange rate stabilization. In the corollary that follows, we show that the policymaker, who desires to increase markups relative to the flexible-price allocation, should appreciate the currency, increase terms of trade, and decrease the output gap.

LEMMA 3. *Under both financial autarky and complete markets, a monetary policy that generates procyclical (countercyclical) markups also gives rise to a countercyclical (procyclical) output gap and a more (less) stable real exchange rate.*

Proof. Lemma 3 establishes a monotonic relationship between the terms of trade, markups, and output. Since the output gap is zero when markups are equal to one, and for any given state of the world higher markups imply lower output, markups above one necessarily imply that output gaps are negative, and the terms of trade and real exchange rate are lower than they would be under flexible prices. ■

To summarize, lemmas 1, 2, and 3 state that the model we consider does not generate any nonstandard dynamics. The impact of monetary policy or productivity shocks is consistent with standard closed or open economy models. With these mechanisms established for simple monetary rules, we are now ready to consider optimal monetary policy.

3.3 Complete Markets and Cooperative Policy

In this section, we examine the optimal monetary policy for cooperative central banks under complete markets. In this setup, nominal rigidity is the only distortion present. Therefore, replicating flexible prices is an optimal policy that also implements the efficient allocation, which is stated more formally in the proposition below.

PROPOSITION 1. *In complete markets, cooperative central banks maximize (34) subject to (35), (36), (37), and (38), where* $\mathbb{1}_{CP} = 1$ *and* $\mathbb{1}_{CM} = 1$*. The indicators for cooperative policy* $\mathbb{1}_{CP}$ *and complete markets* $\mathbb{1}_{CM}$ *are set to one. The solution is*

$$\hat{\mu}_t = 0.$$

The resulting equilibrium allocation exactly coincides with the flexible-price allocation. Mimicking the flexible-price allocation is the optimal policy under cooperation, and corresponds to the social planner allocation.

Proof. See online appendix B.1. ■

In our setup, constant markups imply flexible prices. Moreover, since labor subsidies remove monopolistic distortions, and as complete markets provide full risk sharing, the resulting allocation is efficient. The finding that the flexible-price allocation is the optimal policy under cooperation, and corresponds to the social planner allocation, aligns with Benigno and Benigno (2006) and Corsetti, Dedola, and Leduc's (2010) results for two large economies. Our contribution here is mostly technical.

3.4 Financial Autarky and Cooperative Policy

In this section, we consider the optimal cooperative monetary policy under financial autarky. To our knowledge, we are the first to consider this case for small open economies. Corsetti, Dedola, and Leduc (2010) consider cooperative policy under financial autarky for two open economies. Before comparing our results, we formulate the optimal cooperative policy problem under financial autarky.

PROPOSITION 2. *In financial autarky, cooperative central banks maximize (34) subject to (35), (36), (37), and (38), where* $\mathbb{1}_{CP} = 1$ *and* $\mathbb{1}_{CM} = 0$. *The indicators for cooperative policy* $\mathbb{1}_{CP}$ *and complete markets* $\mathbb{1}_{CM}$ *are set to one and zero, respectively. The solution is given by*

$$\hat{\mu}_t = \frac{G_2}{F_2}\hat{Z}_t, \tag{41}$$

where

$$\begin{aligned} G_2 &= -\alpha(1+\varphi)((1-\alpha)(\eta\sigma-1)+\sigma(\gamma-1)), \\ F_2 &= \varphi+\sigma+\alpha\eta+\alpha\gamma-2\eta\sigma+\eta^2\sigma+\gamma^2\sigma+\alpha^2\eta^2\sigma-2\gamma\sigma-\alpha^2\eta \\ &\quad -2\alpha\eta^2\sigma-2\alpha\eta\gamma\sigma+2\eta\gamma\sigma+2\alpha\eta\sigma \\ &\quad -2\alpha\varphi-2\eta\varphi-2\gamma\varphi+\alpha^2\varphi+\eta^2\varphi+\gamma^2\varphi+\alpha^2\eta^2\varphi+4\alpha\eta\varphi \\ &\quad +2\alpha\gamma\varphi+2\eta\gamma\varphi-2\alpha\eta^2\varphi-2\alpha^2\eta\varphi-2\alpha\eta\gamma\varphi. \end{aligned}$$

The resulting equilibrium allocation differs from the flexible-price allocation.

Proof. See online appendix B.2. ∎

The resulting markup movement reported in equation (41) is a complicated function of openness, trade elasticities, labor disutility, and risk aversion. Although the monetary authority will deviate from replicating the flexible-price allocation in general, there are four specific calibrations where the flexible-price allocation will be optimal for the policymaker. Corollary 2.1 details each of these four cases.

COROLLARY 2.1. *Cooperative central banks under financial autarky implement the flexible-price allocation, whenever any of the four listed conditions below is satisfied:*

(i) under Cole-Obstfeld conditions, when $\sigma = \gamma = \eta = 1$*;*

(ii) under full home bias, $\alpha = 0$*;*

(iii) whenever $(\eta\sigma - 1)(1 - \alpha) + \sigma(\gamma - 1) = 0$*;*

(iv) for an economy with stable terms of trade, when $\gamma \rightarrow \infty$ *or* $\eta \rightarrow \infty$*.*

Proof. Under (i), (ii), and (iii) in equation (41) we have $G_2 = 0$, which implies constant markups and flexible prices. With respect to (iv), we know that as the export elasticity γ or the import elasticity η increase, the numerator F_2 grows faster than G_2. Or, more formally, the following relationships hold as export elasticity increases: $\lim_{\gamma \rightarrow \infty} \frac{G_2}{\gamma} \rightarrow -\alpha(1 + \varphi)\sigma$, $\lim_{\gamma \rightarrow \infty} \frac{F_2}{\gamma^2} \rightarrow \sigma + \varphi$, thus $\lim_{\gamma \rightarrow \infty} \frac{G_2}{F_2} = \hat{\mu}_t \rightarrow 0$. Also, the following relationships apply as import elasticity grows: $\lim_{\eta \rightarrow \infty} \frac{G_2}{\eta} \rightarrow -\alpha(1 - \alpha)(1 + \varphi)\sigma$, $\lim_{\eta \rightarrow \infty} \frac{F_2}{\eta^2} \rightarrow (1 - \alpha)^2(\sigma + \varphi)$, thus $\lim_{\eta \rightarrow \infty} \frac{G_2}{F_2} = \hat{\mu}_t \rightarrow 0$. ■

Under cooperative policy and financial autarky policymakers face two distortions: nominal rigidities and incomplete cross-country risk sharing. Optimal monetary policy in this setting thus faces a trade-off between mitigating the distortionary impact of nominal rigidities through price stability as well as the mitigation of incomplete cross-country risk sharing via terms-of-trade adjustments. Under Cole-Obstfeld conditions, where $\sigma = \gamma = \eta = 1$, terms-of-trade movements provide full risk sharing, such that the optimal monetary policy is to mimic the flexible-price allocation and thereby eliminate distortions from nominal rigidities. In a closed economy

($\alpha = 0$), risk sharing cannot be improved by trade flows, and the central bank thus focuses on eliminating the internal distortion arising from nominal rigidities via replication of the flexible-price allocation. Finally, as the trade elasticities increase (η, γ), monopoly power at the national level is reduced, and policymakers are less able to influence the terms of trade through monetary policy. To summarize, policymakers focus on price stability when terms-of-trade movements provide full risk sharing under flexible prices, or monetary policy is powerless to reduce this risk sharing.

How do cooperative central banks improve risk sharing across countries? Under cooperative monetary policy, countries with positive productivity shocks reduce their markups and depreciate their currencies and terms of trade via lower interest rates in order to supply exports for the rest of the world at reduced prices. On the other hand, countries with negative productivity shocks increase their interest rate, markups, and terms of trade in order to sell exports at higher prices. This cooperative monetary policy response to asymmetric shocks stabilizes employment and consumption across countries.

Corollary 2.2 below summarizes formally the conditions required for countercyclical markups.

COROLLARY 2.2. *If* $\gamma > 1-\eta(1-\alpha)+\frac{1-\alpha}{\sigma}$, *then markup* $\hat{\mu}_t$ *negatively co-moves with output.*

Proof.

$$\hat{\mu}_t = -\alpha \frac{(1-\alpha)(\eta\sigma - 1) + \sigma(\gamma - 1)}{(1-\alpha)^2(\eta - 1)^2 + (\gamma - 1)^2 + 2\eta\gamma(1-\alpha) + 2\alpha\gamma - 1} \hat{Y}_t. \tag{42}$$

Since $\gamma \geq 1$, $\eta > 0$, and $\alpha > 0$, the denominator in equation (42) is positive. Indeed, the denominator monotonically increases with γ. One can also show that for $\gamma = 1$, the denominator is equal to $((1-\alpha)\eta-\alpha)^2 \geq 0$. Therefore, the denominator is positive for $\gamma > 1$. Under $\gamma > 1 - \eta(1-\alpha) + \frac{1-\alpha}{\sigma}$, the numerator in the fraction is positive as well. Therefore, overall coefficient on the right-hand side is negative. ∎

For example, if $\sigma \geq 2$, then $\gamma \geq 1.5$ guarantees that, regardless of other parameters, the policymaker reduces markups in response to higher output, further boosting production. Since in the data the trade elasticity η is positive and risk aversion σ is ≥ 2, markups will co-move negatively with output for almost all parameter values. As a result, cooperative monetary policy in financial autarky yields countercyclical markups and procyclical output gaps, and manipulates the terms of trade to increase cross-border risk sharing.

Corsetti, Dedola, and Leduc (2010) study cooperative policy under financial autarky for two open economies. We differ from their research in several dimensions. First, we use prices set one period in advance instead of Calvo pricing. Second, we allow trade elasticities to be different from each other. Third, they analyze first-order conditions without solving fully for each allocation. For example, they find out that the optimal policy under financial autarky faces a tradeoff between stabilizing output gaps, price dispersion, and global demand imbalances. However, their work does not provide an explicit roadmap for when the central bank should seek to depreciate or appreciate the exchange rate given certain parameter values.

Our analytical solution allows us to see the tradeoff between the output gap stabilization and imperfect risk sharing explicitly. In particular, manipulation of the terms of trade should deliver risk sharing, and the policymakers try to move international prices to deliver higher risk sharing relative to maximizing output gaps. We solve for the explicit policy rule under central bank cooperation: countries with a positive shock lower their interest rate and depreciate their currency and terms of trade relative to the flexible-price allocation in order to provide cheaper products to the rest of the world and stabilize consumption abroad.

Our analytical solution allows us to investigate the role of specific parameters. When trade elasticities are low (η and γ close to one), goods are less substitutable across countries, and central banks exert a stronger influence on the terms of trade because of monopoly power at the export level. However, under low trade elasticities the terms of trade also provide a high degree of risk sharing across countries, mitigating the need for central banks to deviate from the flexible-price allocation. On the other hand, as trade elasticities increase, the terms of trade provide less risk sharing across countries without policy intervention, and thus policymakers will manipulate the terms of trade to improve risk sharing and move the economy closer to the efficient allocation. The caveat is that as goods become more substitutable, policymakers' capacity to influence the terms of trade declines. As a result, under high trade elasticities, central banks will focus more on closing national output gaps and replicating the flexible-price allocation than improving risk sharing. Overall, optimal monetary policy will be closer to the flexible-price allocation for low and high trade elasticities, while deviations from flexible prices will be strongest for intermediate values of the trade elasticities.

3.5 Financial Autarky and Noncooperative Policy: Introduction of the Export Share

In this section, we study noncooperative monetary policy under financial autarky. There are three distortions that may drive the equilibrium away from the first-best allocation. In addition to nominal rigidities and incomplete risk sharing across countries, policymakers exploit terms-of-trade externalities to boost national welfare. Moreover, while incomplete cross-country risk sharing moves the equilibrium away from the first best, this distortion only has an indirect effect on the policymaker, who is disinterested in providing cross-country risk sharing. Nevertheless, market incompleteness affects the dynamics of the endogenous variables and has an effect on the tradeoff the policymaker faces between terms-of-trade externalities and nominal rigidities. In general, the policymaker chooses to deviate from the flexible-price equilibrium in this environment. Proposition 3 below formally establishes that optimal markups deviate from one, such that the flexible-price allocation is suboptimal.

PROPOSITION 3. *In financial autarky, noncooperative central banks maximize (34) subject to (35), (36), (37), and (38), where the indicators for cooperative policy* $\mathbb{1}_{CP}$ *and complete markets* $\mathbb{1}_{CM}$ *are set to zero. The solution is*

$$\hat{\mu}_t = \frac{G_3}{F_3}\hat{Z}_t, \tag{43}$$

where

$$
\begin{aligned}
G_3 &= -\alpha(1-\alpha)(1+\varphi)(\eta-1)(\eta-1+\gamma),\\
F_3 &= 2\alpha\varphi - \sigma - \alpha\gamma - \varphi + \alpha\sigma + 3\eta\varphi + 3\gamma\varphi + 3\eta\sigma + 3\gamma\sigma + \alpha^2\eta\\
&\quad - \alpha^3\eta + \alpha\gamma^2 - \alpha^2\varphi - 3\eta^2\varphi + \eta^3\varphi - 3\gamma^2\varphi + \gamma^3\varphi\\
&\quad - 3\eta^2\sigma + \eta^3\sigma - 3\gamma^2\sigma + \gamma^3\sigma - \alpha^2\eta^2 + \alpha^3\eta^2 - 7\alpha^2\eta^2\varphi\\
&\quad + 3\alpha^2\eta^3\varphi + 2\alpha^3\eta^2\varphi - \alpha^3\eta^3\varphi - 5\alpha^2\eta^2\sigma + 3\alpha^2\eta^3\sigma + \alpha^3\eta^2\sigma\\
&\quad - \alpha^3\eta^3\sigma + \alpha\eta\gamma - 7\alpha\eta\varphi - 4\alpha\gamma\varphi - 5\alpha\eta\sigma - 2\alpha\gamma\sigma - 6\eta\gamma\varphi\\
&\quad - 6\eta\gamma\sigma - \alpha^2\eta\gamma + 8\alpha\eta^2\varphi + 5\alpha^2\eta\varphi - 3\alpha\eta^3\varphi - \alpha^3\eta\varphi\\
&\quad + 2\alpha\gamma^2\varphi + \alpha^2\gamma\varphi + 7\alpha\eta^2\sigma + 2\alpha^2\eta\sigma - 3\alpha\eta^3\sigma + \alpha\gamma^2\sigma + 3\eta\gamma^2\varphi\\
&\quad + 3\eta^2\gamma\varphi + 3\eta\gamma^2\sigma + 3\eta^2\gamma\sigma + 8\alpha\eta\gamma\sigma - 3\alpha\eta\gamma^2\varphi\\
&\quad - 6\alpha\eta^2\gamma\varphi - 4\alpha^2\eta\gamma\varphi - 3\alpha\eta\gamma^2\sigma - 6\alpha\eta^2\gamma\sigma - 2\alpha^2\eta\gamma\sigma\\
&\quad + 3\alpha^2\eta^2\gamma\varphi + 3\alpha^2\eta^2\gamma\sigma + 10\alpha\eta\gamma\varphi.
\end{aligned}
$$

Proof. See online appendix B.3. ∎

Equation (43) describes markup fluctuations as a function of technology shocks. While (43) is more complex than the relationship we can obtain by expressing markup dynamics in terms of the other endogenous variables, it clearly reveals under what conditions the flexible-price allocation is optimal. Corollary 3.1 lists the conditions when the flexible-price allocation is optimal.

COROLLARY 3.1. *Noncooperative central banks under financial autarky implement the flexible-price allocation whenever any of the four listed conditions below is satisfied:*

- *(i) under unitary import elasticity,* $\eta = 1$ *(including Cole-Obstfeld conditions* $\sigma = \gamma = \eta = 1$*);*
- *(ii) under full home bias,* $\alpha = 0$;
- *(iii) under no home bias,* $\alpha = 1$;
- *(iv) under stable terms of trade,* $\gamma \to \infty$ *or* $\eta \to \infty$*;*
- *(v) under extreme risk aversion,* $\sigma \to \infty$.

Proof. Under (i), (ii), and (iii) in equation (43) we have $G_2 = 0$, which implies constant markups and flexible prices. With respect to (iv), as the export elasticity γ increases, the numerator F_3 grows faster than G_3. More formally, $\lim_{\gamma\to\infty} \frac{G_3}{\gamma} \to -\alpha(1-\alpha)(1+\varphi)(\eta-1)$, $\lim_{\gamma\to\infty} \frac{F_3}{\gamma^3} \to \sigma+\varphi$. Thus, markups become more stable as the export elasticity increases and eventually become constant: $\lim_{\gamma\to\infty} \hat{\mu}_t = \lim_{\gamma\to\infty} \frac{G_3}{F_3} = 0$. We also know that as the import elasticity η increases, the numerator F_3 grows faster than G_3. More formally, $\lim_{\eta\to\infty} \frac{G_3}{\eta^2} \to -\alpha(1-\alpha)(1+\varphi)$, $\lim_{\eta\to\infty} \frac{F_3}{\eta^3} \to (1-\alpha)^3(\sigma+\varphi)$. Thus, markups become more stable as the import elasticity grows and eventually become constant: $\lim_{\eta\to\infty} \hat{\mu}_t = \lim_{\eta\to\infty} \frac{G_3}{F_3} = 0$. Finally, as risk aversion $\sigma \to \infty$, the numerator G_3 remains constant, while the denominator $F_3 \to \infty$. Thus, markups become more stable or $\lim_{\sigma\to\infty} \hat{\mu}_t = \lim_{\sigma\to\infty} \frac{G_3}{F_3} = 0$. ■

In financial autarky, the intuition for case (ii) and (iv) in corollary 3.1 is similar under cooperative and noncooperative policy. In a closed economy (case (ii)), manipulation of the terms of trade does not generate any monopolistic rents, as there are no exports. Thus, optimal cooperative and noncooperative policy focuses on replicating the flexible-price allocation. Regarding case (iv), as goods become more substitutable, national monopoly power at the export level is reduced, and in the limit the central bank's ability to influence the terms of trade through monetary policy is eliminated. As a result, optimal monetary policy will focus on alleviating the distortion from nominal rigidities and the central bank will replicate the flexible-price allocation. However, this is where the similarities between optimal cooperative and noncooperative policies in financial autarky end.

Under extreme risk aversion in part (v), the flexible-price allocation stabilizes consumption. Under cooperative policy in financial autarky, joint manipulation of the terms of trade allows policymakers to overcome their inability to share risk through international asset markets via the adjustment of international prices to stabilize consumption across countries. In the absence of coordination, deviation from flexible prices causes consumption and labor to fluctuate, which is infinitely costly under extreme risk aversion. Thus, replication of the flexible-price allocation is optimal for noncooperative policymakers in financial autarky.

Also, the results on cooperative policies are no longer relevant for part (i) and (iii) of corollary 3.1. In a fully open economy and under unitary import elasticity, the flexible-price allocation is optimal. In both cases the share of goods exported is constant and independent from technology shocks. While in a fully open economy the export share equals one, under unitary import elasticities the export share is equal to the degree of openness α. We can formally show this by plugging (38) into (36), where we set $\eta = 1$, $\mathbb{1}_{CM} = 0$, and $\mathbb{1}_{CP} = 0$.

In corollary 3.1 we see that a constant export share implies constant markups, which implies a relationship between the dynamics of the export share and optimal markups. Corollary 3.2 below formally establishes this relationship.

COROLLARY 3.2. *Under the optimal noncooperative monetary policy in financial autarky, there is positive co-movement between the markup and the export share.*

Proof. We can express markup dynamics in terms of export share dynamics using the results from online appendix B.3, equation (B.77):

$$\hat{\mu}_t = \frac{\alpha(\eta + \gamma - 1)}{(\eta(1-\alpha) + \gamma - 1)^2 + \alpha(\eta(1-\alpha) + \gamma - 1)} \hat{E}_{s,t}. \tag{44}$$

For $\gamma \geq 1, \eta > 0, 0 < \alpha < 1$, we have $\frac{\alpha(\eta+\gamma-1)}{(\eta(1-\alpha)+\gamma-1)^2+\alpha(\eta(1-\alpha)+\gamma-1)} > 0$. ∎

Corollary 3.2 holds that when the export share goes up, the markup should increase. Since markups negatively co-move with the output gap, the latter should also negatively correlate with the export share. In other words, monetary policy should be expansionary when the export share goes down, and contractionary when the share of goods exported goes up.

Why do optimal markups tend to move with the share of exports in production? In the steady state our model sets markups equal to one, which allows the implementation of the efficient cooperative steady state. However, under noncooperative policy, the goal of the policymaker is to charge only domestic households markups equal to one. Since the country has monopolistic power with an elasticity of foreign demand equal to γ, the policymakers try to maximize their monopolistic rents by selling goods abroad with markups $\frac{\gamma}{\gamma-1}$. But under the law of one price, the markup is the same regardless of whether producers sell goods abroad or at home. Thus, the optimal markup is an average between one and $\frac{\gamma}{\gamma-1}$. One can show that in the symmetric steady state the optimal noncooperative markup is equal to $1+\frac{\alpha}{\gamma-1+(1-\alpha)\eta}$.

LEMMA 4. *In the steady state, noncooperative social planners maximize (34) subject to (36), (37), and (38), where the indicators for cooperative policy* $\mathbb{1}_{CP}$ *and complete markets* $\mathbb{1}_{CM}$ *are set to zero. The technology level* Z *is equal to one, as there are no technology shocks in steady state. The optimal markup for the social planner in the steady state is*

$$\mu = 1 + \frac{\alpha}{\gamma - 1 + (1-\alpha)\eta}.$$

Proof. See online appendix C.3. ∎

The optimal markup converges to one for the closed economy as α approaches zero, and converges to $\frac{\gamma}{\gamma-1}$ for the fully open economy as α approaches one. In addition, the optimal steady-state markup increases monotonically with α.

LEMMA 5. *The optimal markup for the noncooperative social planner in the steady state monotonically increases with openness:*

$$\frac{\partial \mu}{\partial \alpha} > 0.$$

Proof. $\frac{\partial \mu}{\partial \alpha} = \frac{\gamma+\eta-1}{(\gamma-1+(1-\alpha)\eta)^2} > 0$ for $0 < \alpha < 1$, $\gamma \geq 1$, and $\eta > 0$. ∎

Lemmas 4 and 5 illustrate that steady-state markups increase with openness α, which is equal to the export share in steady state. In our model, we set the steady-state markup to one. However, the optimal markup changes over the cycle since the export share changes. With a higher export share, the optimal monetary policy is to generate higher markups.

The closest relevant study was conducted by De Paoli (2009b), where she considered the small open economy as a limiting case of two large open economies. We differ from her paper in three dimensions. First, she uses Calvo pricing, while we utilize prices set one period in advance. Second, we differentiate between export and import elasticities. Finally, we analyze the solution for the global set of parameters, instead of focusing on a particular calibration. Overall, our results are consistent with De Paoli (2009b), and we confirm in the global parameter-space that the sign of the markup and the output gap depends on whether the import elasticity is greater or less than one. Our main contribution is the presentation of sufficient statistics for optimal policy, which holds whether we take trade elasticity estimates from the macro or trade literature.

3.6 Complete Markets and Noncooperative Policy

Optimal monetary policy for noncooperative central banks under complete markets is the most complex of the four cases we consider. Two distortions drive equilibrium away from the first-best allocation: nominal rigidities and terms-of-trade externalities. The central bank faces a tradeoff between replicating the flexible-price allocation or extracting terms-of-trade rents. The main complicating factor in complete markets is that consumption dynamics differ from output dynamics. The decoupling of consumption from output makes complete markets different from financial autarky or the steady state. As a result, optimal monetary policy is more complicated. Before moving to the intuition, we state the principles for the optimal policy formally.

PROPOSITION 4. *In complete markets, noncooperative central banks maximize (34) subject to (35), (36), (37), and (38), where the indicators for cooperative policy $\mathbb{1}_{CP}$ and complete markets $\mathbb{1}_{CM}$ are set to zero and one, correspondingly. The solution is given by*

$$\hat{\mu}_t = \frac{G_4}{F_4}\hat{Z}_t, \tag{45}$$

where

$$
\begin{aligned}
G_4 &= \sigma\alpha(1-\alpha)(1+\varphi)((1-2\eta)(\eta\sigma-1)\alpha+(\eta-1)^2 \\
&\quad +\eta\sigma(\gamma-1)+\eta^2(\sigma-1)), \\
F_4 &= 4\alpha\varphi-\sigma-\varphi+3\alpha\sigma+\eta\varphi+\gamma\varphi+\eta\sigma+\gamma\sigma-6\alpha^2\varphi+4\alpha^3\varphi \\
&\quad -\alpha^4\varphi-3\alpha^2\sigma+\alpha^3\sigma+2\alpha\eta^2\sigma^2+3\alpha^2\eta\sigma^2-\alpha^3\eta\sigma^2+\alpha\gamma^2\sigma^2 \\
&\quad -5\alpha\eta\varphi-4\alpha\gamma\varphi-5\alpha\eta\sigma-2\alpha\gamma\sigma-5\alpha^2\eta^2\sigma^2+3\alpha^3\eta^2\sigma^2 \\
&\quad +10\alpha^2\eta\varphi-10\alpha^3\eta\varphi+5\alpha^4\eta\varphi-\alpha^5\eta\varphi+6\alpha^2\gamma\varphi-4\alpha^3\gamma\varphi \\
&\quad +\alpha^4\gamma\varphi-2\alpha\eta\sigma^2+7\alpha^2\eta\sigma-3\alpha^3\eta\sigma-\alpha\gamma\sigma^2+\alpha^2\gamma\sigma-2\alpha\eta\varphi\sigma \\
&\quad -2\alpha\gamma\varphi\sigma-\alpha^2\eta^2\varphi\sigma^2+\alpha^2\eta^3\varphi\sigma^2+2\alpha^3\eta^2\varphi\sigma^2-3\alpha^3\eta^3\varphi\sigma^2 \\
&\quad -\alpha^4\eta^2\varphi\sigma^2+3\alpha^4\eta^3\varphi\sigma^2-\alpha^5\eta^3\varphi\sigma^2-\alpha^2\gamma^2\varphi\sigma^2+\alpha^2\gamma^3\varphi\sigma^2 \\
&\quad +3\alpha\eta\gamma\sigma^2+2\alpha\eta^2\varphi\sigma+6\alpha^2\eta\varphi\sigma-6\alpha^3\eta\varphi\sigma+2\alpha^4\eta\varphi\sigma \\
&\quad +2\alpha\gamma^2\varphi\sigma+4\alpha^2\gamma\varphi\sigma-2\alpha^3\gamma\varphi\sigma-3\alpha^2\eta\gamma\sigma^2-8\alpha^2\eta^2\varphi\sigma \\
&\quad +12\alpha^3\eta^2\varphi\sigma-8\alpha^4\eta^2\varphi\sigma+2\alpha^5\eta^2\varphi\sigma-4\alpha^2\gamma^2\varphi\sigma+2\alpha^3\gamma^2\varphi\sigma \\
&\quad -2\alpha^2\eta\gamma\varphi\sigma^2+2\alpha^3\eta\gamma\varphi\sigma^2+4\alpha\eta\gamma\varphi\sigma+3\alpha^2\eta\gamma^2\varphi\sigma^2 \\
&\quad +3\alpha^2\eta^2\gamma\varphi\sigma^2-3\alpha^3\eta\gamma^2\varphi\sigma^2-6\alpha^3\eta^2\gamma\varphi\sigma^2+3\alpha^4\eta^2\gamma\varphi\sigma^2 \\
&\quad -12\alpha^2\eta\gamma\varphi\sigma+12\alpha^3\eta\gamma\varphi\sigma-4\alpha^4\eta\gamma\varphi\sigma.
\end{aligned}
$$

Proof. See online appendix B.4. ■

Equation (45) describes markup dynamics as a function of technology shocks. While it is possible to express the markup in terms of output or other endogenous variables in a more transparent way, equation (45) shows the set of parameters under which the replication of flexible prices is optimal. Corollary 4.1 lists the conditions under which replication of flexible-price allocation is optimal.

COROLLARY 4.1. *Noncooperative central banks under complete markets implement the flexible-price allocation, whenever any of the four listed conditions below is satisfied:*

(i) *under Cole-Obstfeld conditions* ($\sigma = \gamma = \eta = 1$)*;*

(ii) under full home bias ($\alpha = 0$);

(iii) under no home bias ($\alpha = 1$);

(iv) for an economy with stable terms of trade or whenever $\gamma \to \infty$ or $\eta \to \infty$.

Proof. Under (i), (ii), and (iii) in equation (45) we have $G_4 = 0$, which implies constant markups and flexible prices. With respect to (iv), we know that as the export elasticity γ increases, the numerator F_4 grows faster than G_4. More formally, $\lim_{\gamma\to\infty} \frac{G_4}{\gamma} \to \alpha\sigma^2\eta(1-\alpha)(1+\varphi)$, $\lim_{\gamma\to\infty} \frac{F_4}{\gamma^3} \to \alpha^2\varphi\sigma^2$. Thus, markups become more stable as the export elasticity increases: $\lim_{\gamma\to\infty} \hat{\mu}_t = \lim_{\gamma\to\infty} \frac{G_4}{F_4} = 0$. As the import elasticity η increases, the denominator F_4 grows faster than the numerator G_4. More formally, $\lim_{\eta\to\infty} \frac{G_4}{\eta^2} \to -\sigma\alpha(1-\alpha)(1+\varphi)(1+\sigma-2\sigma\alpha)$, $\lim_{\eta\to\infty} \frac{F_4}{\eta^3} \to \alpha^2(1-\alpha)^3\sigma^2\varphi$. Thus, markups become more stable as the import elasticity increases and eventually become constant: $\lim_{\eta\to\infty} \hat{\mu}_t = \lim_{\eta\to\infty} \frac{G_4}{F_4} = 0$. ■

Similar to both cooperative and noncooperative policies under financial autarky, price stability is the optimal noncooperative policy in complete markets when the economy is fully closed (part (ii)) or when the terms of trade are always stable (part (iv)). In a closed economy, manipulation of the terms of trade does not generate any extra benefits because there are no exports, while for stable terms of trade the central bank cannot exploit the terms of trade externality to enhance domestic welfare.

Parts (i) and (iii) are similar to noncooperative policy under financial autarky. Under the Cole-Obstfeld calibration ($\sigma = \eta = \gamma = 1$) or in a fully open economy ($\alpha = 1$), mimicking the flexible-price allocation is optimal. In both cases the export share remains constant. In financial autarky there is a positive relationship between the markup and the export share, but unfortunately the relationship between the two is more complicated in complete markets. Using the solution for the markup from equation (45) and other endogenous variables from online appendix B.4, we obtain the following relationship between the optimal markup and export share:

$$\hat{\mu}_t = \frac{G_x}{F_x H_x} \hat{E}_{s,t}, \tag{46}$$

where

$$\begin{aligned} G_x &= \alpha\sigma(\sigma(1-2\alpha)\eta^2 + \gamma\sigma\eta - (1-\alpha)\eta(\sigma+2) + 1 - \alpha), \\ F_x &= \alpha - 1 + \sigma(\gamma - \alpha\eta), \\ H_x &= (\eta(1-\alpha) + \gamma - 1)(1-\alpha)^2 \\ &\quad + \alpha\sigma[\eta(\eta(1-\alpha) + 2(2\gamma - 1))(1-\alpha) + \gamma(\gamma - 1)]. \end{aligned}$$

In (46) the relationship between the optimal markup and the export share is no longer monotonic. The key intuitive difference between financial autarky and complete markets is that the former behaves more like the steady state in comparative statics. The monetary policy tradeoff under complete markets is to adjust the markup to extract higher terms-of-trade rents without generating too much volatility in consumption and hours worked. If consumption and labor become too volatile when the central bank attempts to exploit terms-of-trade externalities, it may be optimal for the policymaker to move toward the flexible-price allocation and away from terms-of-trade manipulation. However, the more extracting higher terms-of-trade rents actually reduces the volatility of consumption and labor, the more optimal policy will deviate from replicating flexible prices.

As the parameter governing labor disutility is absent in (46), we focus now on the volatility of consumption. Consider a positive productivity shock. In general, under flexible prices the terms of trade should go down, consumption should go up, and the export share may increase or decrease. If the export share increases following a positive productivity shock, the terms-of-trade externality creates upward pressure on the markup, such that the policymaker will appreciate the terms of trade and the real exchange rate and decrease the level of consumption. Thus, a higher markup allows the policymaker to stabilize consumption and extract higher monopolistic rents, such that an increase in the markup is likely to be a dominant strategy when the export share is positively correlated with technology. On the other hand, if the export share declines after a positive productivity shock, the policymaker faces a different tradeoff. In this case, the terms-of-trade externality creates downward pressure on the markup. However, a lower markup will depreciate the terms of trade and the real exchange rate, and increase consumption. Since consumption increases on impact, a further increase is undesirable when households are risk averse due to higher consumption volatility. Corollary 4.2 below formally establishes the relationship we have described.

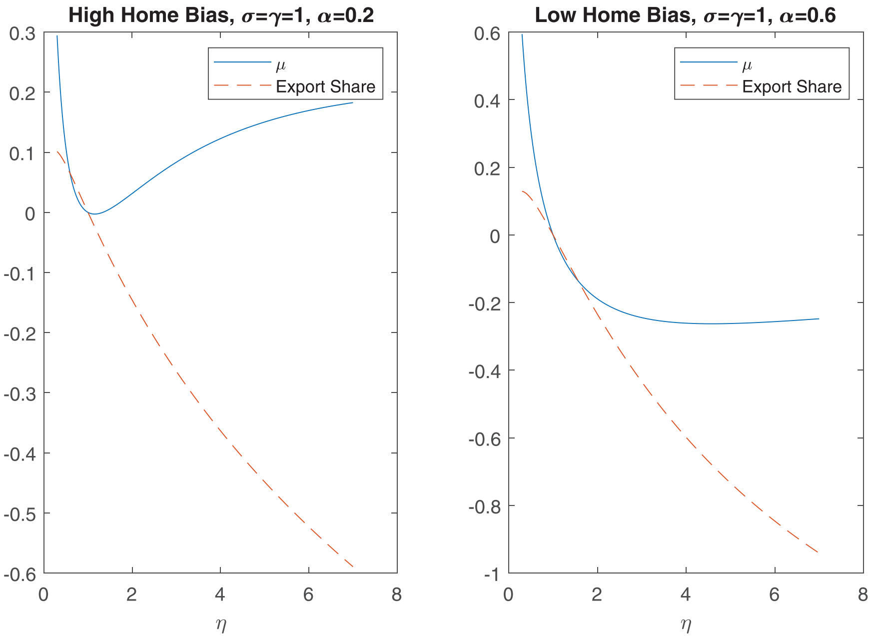
COROLLARY 4.2. *Under complete markets and noncooperative policy, markups are procyclical if the export share is procyclical and risk aversion (σ) is greater than one.*

Proof. See online appendix C.4. ∎

Under what conditions might the export share decline after a positive productivity shock? First, this might happen if domestic consumption goes up too strongly in response to a shock. For example, under complete markets, consumption is more sensitive to the real exchange rate for lower values of risk aversion, and a positive productivity shock typically leads to a depreciation in the terms of trade and the real exchange rate and a rise in consumption. The sensitivity of consumption to the real exchange is inversely proportional to the coefficient of risk aversion. However, while a lower degree of risk aversion leads to higher consumption volatility, it also decreases the welfare loss from increased volatility. The second factor is home bias. The real exchange rate becomes more sensitive to the terms of trade as home bias increases. Thus, for a given fall in the terms of trade, the real exchange rate will depreciate and consumption will rise more strongly as home bias increases. The third factor is the relative magnitude of the import and export elasticities. If the import elasticity is significantly larger than the export elasticity, a decrease in the terms of trade will generate a rise in demand for the home good from domestic consumers that outstrips the rise in demand from abroad, leading to a decline in the export share.

To generate a disconnect between the export share and the markup, we need $\eta > \gamma$ in combination with a low degree of consumption home bias. In figure 2 we deviate from the Cole-Obstfeld

Figure 2. Response in Percent of the Markup and Export Share to a 1 Percent Increase in Technology Level



calibration by increasing η and considering both a low and high degree of home bias.

Figure 2 plots the reaction of the markup and the export share for a particular combination of parameters following a 1 percent increase in technology. For example, the value of 0.2 on the vertical axis reflects that the markup or export share goes up by 0.2 percent in response to a 1 percent increase in the technology level. The subplot on the left corresponds to high home bias ($\alpha = 0.2$) while the subplot on the right corresponds to low home bias ($\alpha = 0.6$). Relative risk aversion and the elasticity of export substitution equal one in both subplots. When the import substitution elasticity η is equal to one, we return to the Cole-Obstfeld case, where the export share is constant for any realization of the technology shock and the flexible-price allocation is optimal. Thus, under the Cole-Obstfeld calibration, the markup and export share lines intersect at zero on the vertical axis and $\eta = 1$ on the horizontal axis as both the markup and the export share are constant in this case. Under high home bias and high elasticity of import substitution η, the markup responds positively to a technology shock, while the export share declines. On the other hand, for low home bias, the markup and export share decline in response to a positive shock when home and foreign goods are substitutes ($\eta > 1$), while they both increase after a positive technology shock when home and foreign goods are complements ($\eta < 1$).

The co-movement between the export share and markup is broken for high home bias and high η, but not for low home bias or low η. Why is this the case? With a positive technology shock, high home bias, and high η, the export share shrinks, and the terms of trade externality creates an incentive to reduce the markup. However, lowering the terms of trade in order to reduce markups has a strong effect on the real exchange rate and causes a substantial consumption increase when consumption is already high. This strong pressure from excess consumption volatility causes the policymaker to push markups in the opposite direction to stabilize consumption. Under low home bias, a reduction in the terms of trade has a small effect on the real exchange rate and causes only a small increase in consumption. In this case, markups can be countercyclical when the export share is countercyclical.

To summarize our findings on noncooperative policy in complete markets, under the most realistic calibrations the optimal markup and the export share are procyclical. Monetary policy will generate negative output gaps in response to positive technology shocks by appreciating the real exchange rate and the terms of trade. We prove that when the export share is procyclical, optimal monetary policy requires the central bank engineer procyclical markups and countercyclical output gaps.

The closest relevant study was conducted by De Paoli (2009a, 2009b), who considered the small open economy as a limiting case of two open economies. As in the previous section, we differ from her paper in three dimensions. First, she uses Calvo pricing, while we utilize prices set one period in advance. Second, we differentiate between export and import elasticities. Finally, we analyze the solution for a global set of parameters, instead of focusing on a particular calibration.

For noncooperative policy under complete markets, we find that differentiation between import and export elasticities plays a major role. Corollary 4.3 below shows that the markup and the export share are positively correlated.

COROLLARY 4.3. *If the import substitution elasticity η is equal to the export substitution elasticity γ, and $\gamma > 1$, then the markup and the export share co-move positively.*

Proof. See online appendix C.5. ■

Whenever relative risk aversion is greater than one, the co-movement between the markup and the export share disappears only if the import elasticity is greater than the export elasticity.

4. Conclusion

There is a long tradition in macroeconomics of explaining the data by introducing distortions into perfectly competitive and efficient markets. In the closed economy this strategy has been fruitful and brought some immediate results in the form of the divine coincidence: by stabilizing inflation the central bank eliminates distortions arising from nominal rigidities, closes the output gap, and removes price dispersion. Such a strategy has been less successful in the context of open economies, where imperfect cross-country risk sharing and terms-of-trade externalities in addition to nominal rigidities prevent the divine coincidence.

Our aim here is to establish some simple principles for optimal cooperative and noncooperative monetary policy in small open economies under complete markets and financial autarky. Relative to the literature, we do not consider the small open economy as the limiting case of two large economies, which allows us to differentiate between the export and import elasticity of substitution.

We are the first to consider cooperative optimal monetary policy for small open economies, which enables a clearer understanding of distortions in the absence of strategic interactions between countries. We find that mimicking the flexible-price allocation is the optimal cooperative monetary policy under complete markets, which aligns with studies focused on two large economies. We also establish that under most realistic calibrations, cooperative optimal monetary policy under financial autarky deviates from price stability in favor of countercyclical markups, procyclical output gaps, and volatile terms of trade.

We then examine noncooperative policy. Under financial autarky, markups set up by the monetary authority co-move positively with the share of the goods exported regardless of the degree of risk aversion, home bias, or product substitutability across countries. Under complete markets, optimal noncooperative monetary policy should generate procyclical markups and countercyclical output gaps whenever the export share is procyclical. This rule may be violated if consumption is strongly sensitive to monetary policy and is negatively correlated with the share of goods exported. In this case, central banks should restrain from lowering markups during a boom even if the export share falls, since responding to such movements with a lower markup might cause excess consumption volatility.

Across all four cases examined, the simple prescription for optimal monetary policy is to replicate flexible prices unless the export share is too volatile. If the export share is procyclical, then the optimal markup should also be procyclical. If the export share is countercyclical, then the optimal markup is countercyclical unless it leads to high consumption volatility.

To conclude, we find a simple monetary policy rule in the noncooperative case. Policymakers should set markups to react to the export share unless a deviation of the markup from zero causes excess consumption volatility.

References

Amiti, M., O. Itskhoki, and J. Konings. 2014. “Importers, Exporters, and Exchange Rate Disconnect.” *American Economic Review* 104 (7): 1942–78.

Anderson, J. E., and E. van Wincoop. 2004. “Trade Costs.” *Journal of Economic Literature* 42 (3): 691–751.

Backus, D. K., P. J. Kehoe, and F. E. Kydland. 1994. “Dynamics of the Trade Balance and the Terms of Trade: The J-curve?” *American Economic Review* 84 (1): 84–103.

Benigno, G., and P. Benigno. 2003. “Price Stability in Open Economies.” *Review of Economic Studies* 70 (4): 743–64.

———. 2006. "Designing Targeting Rules for International Monetary Policy Cooperation." *Journal of Monetary Economics* 53 (3): 473–506.

Benigno, P., and M. Woodford. 2012. "Linear-Quadratic Approximation of Optimal Policy Problems." *Journal of Economic Theory* 147 (1): 1–42.

Cole, H. L., and M. Obstfeld. 1991. "Commodity Trade and International Risk Sharing: How Much Do Financial Markets Matter?" *Journal of Monetary Economics* 28 (1): 3–24.

Corsetti, G., L. Dedola, and S. Leduc. 2010. "Optimal Monetary Policy in Open Economies." In *Handbook of Monetary Economics*, Vol. 3, ed. B. M. Friedman and M. Woodford, 861–933 (chapter 16). Elsevier.

Corsetti, G., and P. Pesenti. 2001. "Welfare and Macroeconomic Interdependence." *Quarterly Journal of Economics* 116 (2): 421–45.

———. 2005. "International Dimensions of Optimal Monetary Policy." *Journal of Monetary Economics* 52 (2): 281–305.

De Paoli, B. 2009a. "Monetary Policy and Welfare in a Small Open Economy." *Journal of International Economics* 77 (1): 11–22.

———. 2009b. "Monetary Policy under Alternative Asset Market Structures: The Case of a Small Open Economy." *Journal of Money, Credit and Banking* 41 (7): 1301–30.

Devereux, M. B., and C. Engel. 2003. "Monetary Policy in the Open Economy Revisited: Price Setting and Exchange-Rate Flexibility." *Review of Economic Studies* 70 (4): 765–83.

Dmitriev, M., and J. Hoddenbagh. 2019. "Optimal Fiscal Transfers in a Monetary Union." *Journal of International Economics* 117 (March): 91–108.

Egorov, K., and D. Mukhin. 2019. "Optimal Monetary Policy under Dollar Pricing." Technical Report.

Faia, E., and T. Monacelli. 2008. "Optimal Monetary Policy in a Small Open Economy with Home Bias." *Journal of Money, Credit and Banking* 40 (4): 721–50.

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

Galí, J., and T. Monacelli. 2005. "Monetary Policy and Exchange Rate Volatility in a Small Open Economy." *Review of Economic Studies* 72 (3): 707–34.

———. 2008. "Optimal Monetary and Fiscal Policy in a Currency Union." *Journal of International Economics* 76 (1): 116–32.

Goldberg, L. S., and C. Tille. 2008. "Vehicle Currency Use in International Trade." *Journal of International Economics* 76 (2): 177–92.

Gopinath, G., E. Boz, C. Casas, F. J. Díez, P.-O. Gourinchas, and M. Plagborg-Møller. 2016. "Dominant Currency Paradigm." NBER Working Paper No. 22943.

Gopinath, G., O. Itskhoki, and R. Rigobon. 2010. "Currency Choice and Exchange Rate Pass-Through." *American Economic Review* 100 (1): 304–36.

Imbs, J., and I. Mejean. 2015. "Elasticity Optimism." *American Economic Journal: Macroeconomics* 7 (3): 43–83.

Kim, J., and S. H. Kim. 2003. "Spurious Welfare Reversals in International Business Cycle Models." *Journal of International Economics* 60 (2): 471–500.

Mundell, R. A. 1963. "Capital Mobility and Stabilization Policy under Fixed and Flexible Exchange Rates." *Canadian Journal of Economics and Political Science/Revue canadienne d'Economique et de Science politique* 29 (4): 475–85.

Obstfeld, M., and K. Rogoff. 1995. "Exchange Rate Dynamics Redux." *Journal of Political Economy* 103 (3): 624–60.

———. 2000. "New Directions for Stochastic Open Economy Models." *Journal of International Economics* 50 (1): 117–53.

———. 2002. "Global Implications of Self-Oriented National Monetary Rules." *Quarterly Journal of Economics* 117 (2): 503–35.

Rogoff, K. 1985. "Can International Monetary Policy Cooperation Be Counterproductive?" *Journal of International Economics* 18 (3–4): 199–217.

Simonovska, I., and M. E. Waugh. 2014a. "The Elasticity of Trade: Estimates and Evidence." *Journal of International Economics* 92 (1): 34 – 50.

———. 2014b. "Trade Models, Trade Elasticities, and the Gains from Trade." NBER Working Paper No. 20495.

Stockman, A. C., and L. L. Tesar. 1995. "Tastes and Technology in a Two-Country Model of the Business Cycle: Explaining International Comovements." *American Economic Review* 85 (1): 168–85.