

Can Macroeconomists Forecast Risk? Event-Based Evidence from the Euro-Area SPF

Geoff Kenny, Thomas Kostka, and Federico Masera

Monetary Policy Committees: Comparing Theory and “Inside” Information from MPC Members

Mikael Apel, Carl Andreas Claussen, Petra Lennartsdotter, and Øistein Røisland

The ECB Unconventional Monetary Policies: Have They Lowered Market Borrowing Costs for Banks and Governments?

Urszula Szczerbowicz

Liquidity Requirements: A Double-Edged Sword

Philipp Johann König

How Much Does the Public Know about the ECB's Monetary Policy? Evidence from a Survey of Dutch Households

Carin van der Cruijssen, David-Jan Jensen, and Jakob de Haan

Model-Based Ex Post Evaluation of Monetary Policy

Eyal Argov, Alon Binyamini, Eliezer Borenstein, and Irit Rozenshtrom

Interest Rate Determination in China: Past, Present, and Future

Dong He, Honglin Wang, and Xiangrong Yu

How Frequently Should We Reestimate DSGE Models?

Marcin Kolasa and Michał Rubaszek



Can Macroeconomists Forecast Risk? Event-Based Evidence from the Euro-Area SPF <i>Geoff Kenny, Thomas Kostka, and Federico Masera</i>	1
Monetary Policy Committees: Comparing Theory and “Inside” Information from MPC Members <i>Mikael Apel, Carl Andreas Claussen, Petra Lennartsdotter, and Øistein Røisland</i>	47
The ECB Unconventional Monetary Policies: Have They Lowered Market Borrowing Costs for Banks and Governments? <i>Urszula Szczerbowicz</i>	91
Liquidity Requirements: A Double-Edged Sword <i>Philipp Johann König</i>	129
How Much Does the Public Know about the ECB’s Monetary Policy? Evidence from a Survey of Dutch Households <i>Carin van der Cruijssen, David-Jan Jensen, and Jakob de Haan</i>	169
Model-Based Ex Post Evaluation of Monetary Policy <i>Eyal Argov, Alon Binyamini, Eliezer Borenstein, and Irit Rozenshtrom</i>	219
Interest Rate Determination in China: Past, Present, and Future <i>Dong He, Honglin Wang, and Xiangrong Yu</i>	255
How Frequently Should We Reestimate DSGE Models? <i>Marcin Kolasa and Michał Rubaszek</i>	279

Copyright © 2015 by the Association of the International Journal of Central Banking.
All rights reserved. Brief excerpts may be reproduced or translated provided the source
is cited. Consult www.ijcb.org for further information.

The *International Journal of Central Banking* is published quarterly
(ISSN: 1815-4654). Online access to the publication is available free of charge
at **www.ijcb.org**. Individual print subscriptions are available. Orders may be placed
by phone (001 415 974 2035), via fax (001 415 974 2168), or by e-mail (editor@ijcb.org).

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

International Journal of Central Banking
Economic Research Department
Federal Reserve Bank of San Francisco
101 Market Street
San Francisco, CA 94105
USA

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

The views expressed in this journal do not necessarily represent the views of the
Association of the International Journal of Central Banking or any of its members.

ISSN: 1815-4654

International Journal of Central Banking

Board of Directors

Chairman

Claudio Borio, *Bank for International Settlements*

Board Members

Q. Farooq Akram, *Norges Bank*
Abdulrahman Al-Hamidy, *Saudi Arabian Monetary Agency*
David E. Altig, *Federal Reserve Bank of Atlanta*
Carlos Hamilton Vasconcelos Araujo, *Central Bank of Brazil*
Jan Marc Berk, *The Nederlandsche Bank*
Mohamed Tahar Bouhouche, *Bank of Algeria*
Lillian Cheung, *Hong Kong Monetary Authority*
Laurent Clerc, *Bank of France*
Francisco G. Dakila Jr., *Central Bank of the Philippines*
Spencer Dale, *Bank of England*
Michael Dotsey, *Federal Reserve Bank of Philadelphia*
William English, *Federal Reserve Bank*
Jiang Feng, *People's Bank of China*
Manuel Ramos Francia, *Bank of Mexico*
Jeffrey C. Fuhrer, *Federal Reserve Bank of Boston*
Kamil Galuscak, *Czech National Bank*
Andrés González Gómez, *Banco de la República*
Niels Lynggård Hansen, *Danmarks Nationalbank*
Philipp Hartmann, *European Central Bank*
Heinz Herrmann, *Deutsche Bundesbank*
Tor Jacobson, *Sveriges Riksbank*
Juan F. Jimeno, *Bank of Spain*
George A. Kahn, *Federal Reserve Bank of Kansas City*
Ali Hakan Kara, *Central Bank of Turkey*
Christopher Kent, *Reserve Bank of Australia*
Evan Koenig, *Federal Reserve Bank of Dallas*
Ana Christina Leal, *Bank of Portugal*

Carlos Lenz, *Swiss National Bank*
Choy Keen Meng, *Monetary Authority of Singapore*
John McDermott, *Reserve Bank of New Zealand*
Maurice McGuire, *Central Bank of Ireland*
Deepak Mohanty, *Reserve Bank of India*
Vyacheslav Morgunov, *Central Bank of Russian Federation*
Alberto Naudon, *Central Bank of Chile*
Edward Offenbacher, *Bank of Israel*
Fabio Panetta, *Bank of Italy*
Thórarinn G. Pétursson, *Central Bank of Iceland*
Cristian Popa, *National Bank of Romania*
Ivan Ribnikar, *Bank of Slovenia*
Lawrence Schembri, *Bank of Canada*
Sam Schulhofer-Wohl, *Federal Reserve Bank of Minneapolis*
Mark Schweitzer, *Federal Reserve Bank of Cleveland*
Jan Smets, *National Bank of Belgium*
Young Kyung Suh, *Bank of Korea*
Daniel Sullivan, *Federal Reserve Bank of Chicago*
Juha Tarkka, *Bank of Finland*
George Tavlas, *Bank of Greece*
Joseph Tracy, *Federal Reserve Bank of New York*
Dobieslaw Tymoczko, *National Bank of Poland*
Christopher Waller, *Federal Reserve Bank of St. Louis*
Kenichirou Watanabe, *Bank of Japan*
John Weinberg, *Federal Reserve Bank of Richmond*

Editorial Board

Managing Editor

John C. Williams

Federal Reserve Bank of San Francisco

Co-editors

Pierpaolo Benigno LUISS Guido Carli	Harrison Hong Princeton University	Rafael Repullo CEMFI
Michael B. Devereux University of British Columbia	Loretta Mester Federal Reserve Bank of Cleveland	Stephanie Schmitt-Grohe Columbia University

Associate Editors

Patrick Bolton University of Columbia	Jordi Galí Centre de Recerca en Economia Internacional (CREI)	Eli M. Remolona Bank for International Settlements
Michael D. Bordo Rutgers University	Marvin Goodfriend Carnegie Mellon University	Hélène Rey London Business School
Mark Carey Federal Reserve Board	Michael B. Gordy Federal Reserve Board	Jean-Charles Rochet University of Toulouse
Pierre Collin-Dufresne University of Columbia	Luigi Guiso European University Institute	Andrew K. Rose University of California, Berkeley
Guy Debelle Reserve Bank of Australia	Andrew G. Haldane Bank of England	Klaus Schmidt-Hebbel Organisation for Economic Co-operation and Development (OECD)
Douglas W. Diamond University of Chicago Graduate School of Business	Takatoshi Ito University of Tokyo	Lars E.O. Svensson Sveriges Riksbank
Francis Diebold University of Pennsylvania	David Lando Copenhagen Business School	Jürgen von Hagen University of Bonn
Michael Dotsey Federal Reserve Bank of Philadelphia	Philip Lane Trinity College Dublin	Ernst-Ludwig von Thadden University of Mannheim
Darrell Duffie Stanford University	Francesco Lippi University of Sassari	Tsutomu Watanabe University of Tokyo
	Carmen M. Reinhart Harvard Kennedy School	

Advisory Board

Franklin Allen The Wharton School of the University of Pennsylvania	Hyun Shin Princeton University	Michael Woodford Columbia University
Charles Goodhart London School of Economics	Kazuo Ueda University of Tokyo	John Taylor Stanford University
	Carl E. Walsh University of California	

Can Macroeconomists Forecast Risk?

Event-Based Evidence from the Euro-Area SPF*

Geoff Kenny,^a Thomas Kostka,^a and Federico Masera^b

^aEuropean Central Bank

^bUniversidad Carlos III de Madrid

We apply methods to evaluate the risk assessments collected as part of the ECB Survey of Professional Forecasters (SPF). Our approach focuses on direction-of-change predictions as well as the prediction of more specific high and low macroeconomic outcomes located in the upper and lower regions of the predictive densities. For inflation and GDP growth, we find such surveyed densities are informative about future direction of change. Regarding high and low outcome events, the surveys are most informative about GDP growth outcomes and at short horizons. The upper and lower regions of the predictive densities for inflation appear less informative.

JEL Codes: C22, C53.

*The authors would like to thank Domenico Giannone, Kajal Lahiri, Tom Stark, Simon van Norden, Shaun Vahey, Jonathan Wright, and Ken Wallis for useful comments on earlier drafts of this paper as well as participants at seminars in the Federal Reserve Bank of Philadelphia (June 29, 2012), the 32nd International Symposium on Forecasting (Boston, June 24–28, 2012), the ECB Workshop on SPF-related research on (November 26, 2012), the joint JCER-ESRI International Conference (Tokyo, February 21, 2013), and the 9th Annual joint Cirano-Cireq Workshop on data revisions in economic forecasting and policy (Montreal, October 10–11, 2013). We are also particularly grateful for many helpful comments and suggestions from three anonymous referees. Any errors are the sole responsibility of the authors. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the ECB or the Eurosystem. Corresponding author (Kenny): DG Research, European Central Bank, Kaiserstrasse 29, D-60311 Frankfurt, Germany. E-mail: Geoff.kenny@ecb.europa.eu.

1. Introduction

Recent experience with macroeconomic forecasting in an environment characterized by high levels of macroeconomic volatility has both highlighted the strong limitations to point forecasts as a sufficient basis for forward-looking policy deliberations *and* strengthened the demand for quality information on the risks surrounding the economic outlook. Indeed, information from the entire predictive densities of future macroeconomic outcomes has an important theoretical justification in the decision sciences (see, for example, Tay and Wallis 2000). Fortunately, such information is increasingly available from different sources and often features in public discussions of the economic outlook. One such source is the Survey of Professional Forecasters (SPF) conducted by the European Central Bank (ECB) on a quarterly basis since the launch of the single currency in January 1999. The Federal Reserve Bank of Philadelphia has an even longer tradition of collecting information on macroeconomists' assessments of future risks via its SPF. Indeed, a large literature has developed around the density forecasts from the U.S. SPF—see, for example, Diebold, Tay and Wallis (1999) Clements (2006, 2010), Giordani and Söderlind (2006), Lahiri and Wang (2007), and Engelberg, Manski, and Williams (2009), among others. In broad terms, these studies highlight the usefulness of the information contained in expert density forecasts—particularly at short horizons—whilst at the same time identifying notable shortcomings, including inattentiveness and overconfidence of forecasters when updating and reporting their assessments.

In this study we adopt a perspective similar to Clements (2006) and examine whether the SPF density forecasts from the ECB SPF provide insights about key future macroeconomic events. To do so, we partition the density at a given threshold point and consider only the binary set of mutually exclusive outcomes, i.e., either the outcome is above the specified threshold or it is below it. We focus on three key broad macroeconomic events: the probability of (i) a relatively low outcome of below 1 percent for the target variables (growth and inflation), (ii) a relatively high outturn of above 2 percent, and finally, (iii) an increase in the forecast target variable compared with its current level. By focusing on expert risk assessments of these “events,” our analysis can reveal aspects of the densities

which are informative *even if* the overall density forecast exhibits a weak performance. This might arise, for example, as a result of heterogeneity in forecasters' loss functions which may lead them to be particularly adept (or inadept) at providing information on the likelihood of particular events. In helping identify aspects of the density forecasts which may be most reliable, our analysis is therefore of primary interest to density forecast users, including both monetary policymakers and those charged with maintaining financial stability.

Our evaluation is based on two key features of the densities. The first is a measure of their *calibration*, which refers to the correspondence between the predicted probabilities and the average frequency of occurrence of the event in question. The second component refers to their *resolution*, which is the main determinant of the densities' usefulness for decision making, as it measures their ability to discriminate between times when the risk materializes and times when it does not.¹ Our analysis is conducted using the aggregate headline SPF density forecasts that are regularly communicated to the public. However, we also pool density forecasts at the individual level and test for the presence of heterogeneity in forecaster skill at the individual level. To do so, we apply a GMM estimation procedure that allows for possible heterogeneity across forecasters as well as aggregate shocks impacting all forecasting agents jointly. The layout of the remainder of the paper is as follows. In section 2, we describe in more detail the evaluation framework we adopt. In section 3, we provide some background descriptive statistics on the events and SPF probability forecasts. Section 4 presents our main empirical findings, while section 5 concludes.

2. Evaluating Event Forecasts

In contrast to point forecasts, a probability forecast for a particular event can never be said to have been either right or wrong, because

¹These two features are based on the decomposition of the quadratic probability score due to Murphy (1973). In an economic context, the Murphy decomposition has been used to evaluate probabilistic forecasts by Diebold and Rudebusch (1989), Lahiri and Wang (2007), and Galbraith and van Norden (2012) but has been much more widely and frequently applied in the statistical and meteorological forecasting literature (Murphy 1988 and Murphy and Winkler 1992). Mitchell and Wallis (2011) also discuss tests of density forecast calibration.

we never observe the “true” probability. However, when such forecasts are issued over a period of time, it is nonetheless possible to apply checks of their “external validity,” i.e., evaluating their correspondence with the related outcome over time. As reviewed in Dawid (1984), a long tradition exists of testing the external validity of probability forecasts in the statistical and meteorological forecasting literature (e.g., Murphy 1973, Yates 1982, Murphy and Winkler 1992, and Gneiting, Balabaoui, and Raftery 2007). Such methods involve gauging the usefulness of such forecasts with respect to the observed outcome. They have been applied to economic forecasting by, among others, Berkowitz (2001), Clements (2006), Lahiri and Wang (2007), Boero, Smith, and Wallis (2011), Mitchell and Wallis (2011), and Galbraith and van Norden (2012) and are closely related to the field of interval forecasting discussed in Christoffersen (1998).² The approach we adopt here is also very much in the spirit of Berkowitz (2001) insofar as we emphasize the evaluation of the entire distribution. However, in contrast to Berkowitz, our approach focuses on the decomposition of the quadratic probability score (QPS), which is a mean squared error (MSE) type scoring function applied to probability forecasts due to Brier (1950). The latter is directly analogous to the MSE of a point forecast, with the exception that the outcome variable ($x_{t+\tau}$) is a binary random variable taking a value of unity when the event occurs and zero if it does not. The $QPS(f_{t+\tau}, x_{t+\tau}) = E[f_{t+\tau} - x_{t+\tau}]^2$ thus provides a scoring rule which penalizes probability forecasts ($f_{t+\tau}$) which are low (high) when the event occurs (does not occur). To shed light on the attributes and validity of probability forecasts, Murphy (1973) suggested a factorization of the QPS based on the conditional distribution of $x_{t+\tau}$ given $f_{t+\tau}$, i.e.,

$$QPS(f_{t+\tau}, x_{t+\tau}) = \sigma_x^2 + E_f[\mu_{x/f} - f]^2 - E_f[\mu_{x/f} - \mu_x]^2. \quad (1)$$

The first term on the right-hand side of (1) measures the unconditional variance of the binary outcome variable, which can be seen

²Granger and Pesaran (2000) argue in favor of a closer link between decisions of forecast users and the forecast evaluation problem, stressing also the importance of predictive distributions. In this respect, the recent work of Andrade, Gyhsels, and Idier (2011) highlights the value in SPF distributions by helping to identify their potential role in the central bank reaction function.

as a proxy for the difficulty of the specific forecasting situation. The second term measures the overall reliability or *calibration* error of the forecasts as the difference between the forecast probability (f) and the expected frequency of occurrence given the forecasts ($\mu_{x|f}$). Well-calibrated probability forecasts are approximately valid or “unbiased in the large” (Murphy and Epstein 1967). All other things equal, miscalibrated forecasts will tend to have a larger QPS. However, even perfectly calibrated forecasts can be clearly unsatisfactory if the forecaster is unable to gauge the timing of the event. The last term on the right-hand side of the equation provides a measure of the *resolution* of the forecasts. Resolution contributes negatively to the QPS, all other things equal. It captures the ability of forecasters to use their probability forecasts to sort individual outcomes into groups which differ from the long run or unconditional relative frequency of occurrence (μ_x). Probability forecasts with high resolution will therefore take values that are further away from the mean frequency of occurrence and closer to the zero or one extremes. The art of probability forecasting can thus be viewed as trying to minimize (1) by optimally managing such a trade-off between the information gain that emerges from having high resolution and the associated reduction in overall accuracy (and miscalibration) that high-resolution forecasts may ultimately introduce.³ Of course, the extent of this trade-off will most likely differ depending on the forecasting situation, e.g., depending on the economic variable, the forecast horizon, or the particular economic context. In the remainder of this section, we describe how we draw inference on these key properties of the SPF probabilities.

2.1 Aggregate-Level Analysis

Murphy and Winkler (1992), Galbraith and van Norden (2012), and Lahiri and Wang (2006, 2007) discuss econometric regression based tests of “perfect” calibration and “zero” resolution (i.e., no skill of forecasters in sorting outcomes). Such tests are based on a generalization of the forecast-realization regressions originally suggested in Mincer and Zarnowitz (1969) to probability forecasts. At

³See Lahiri and Wang (2007) and Galbraith and van Norden (2012) for further discussion of the Murphy decomposition.

the aggregate level, for a given forecast horizon (τ), both tests can be constructed by regression of the outcome in period $t + \tau$ on a constant and the probability forecasts for the same period:

$$x_{t+\tau} = \alpha + \beta f_{t+\tau} + \varepsilon_{i,t+\tau}. \quad (2)$$

As shown by Murphy and Winkler (1992), the fitted values of (2) provide an estimate of the expected probability of the outcome conditional on observing the forecast, i.e., $\mu_{x|f}$ in the decomposition of the QPS in (1). The estimate of forecast calibration, $E_f[\mu_{x|f} - f]^2$, is then computed as the sample averaged (i.e., mean) squared difference between the fitted value of this regression and the probability forecast. Likewise, an estimate of the resolution ($E_f[\mu_{x|f} - \mu_x]^2$) component in (1) corresponds to the mean squared difference between $\mu_{x|f}$ and an estimate for the unconditional mean of the outcome variable of the binary event (μ_x).⁴

As discussed in Murphy and Winkler (1992) and Galbraith and van Norden (2012), the hypothesis of perfectly calibrated forecasts requires that the forecasts do not systematically deviate from $\mu_{x|f}$, which is implied by the restrictions $\alpha = 0$ and $\beta = 1$. As it implies a joint restriction on the model's two parameters, the test of perfect calibration can be implemented as a Wald test using the χ^2 distribution with two degrees of freedom. As discussed in Holden and Peel (1990), albeit in a context that related to point forecasts, this test constitutes a sufficient condition for well-calibrated forecasts that also implies a significant degree of resolution; that is, under the null hypothesis ($\alpha = 0$ and $\beta = 1$), the outcome variable is positively correlated with the forecast. Moreover, joint hypotheses tests of this type may suffer from low power, especially in small samples. For these reasons, Holden and Peel (1990) suggest a necessary and therefore less demanding condition for calibration, namely the direct analysis of the "forecast error." Following this logic, we

⁴In estimating equation (1) we consider the simple discretized version originally suggested in Murphy and Winkler (1992) and Galbraith and van Norden (2012). In contrast, Galbraith and van Norden (2011) apply a non-parametric kernel regression of x on f . The advantage of such a continuous function is that it allows evaluation of calibration at any point in the continuous interval $[0, 1]$. In the present case where the forecast distribution is condensed to only two outcome bins, the two regressions would render the exact same results. Hence, a single regression of the binary outcome variable suffices in our setting.

also compute an additional test of calibration and test directly for a zero mean in the probability forecast error ($x_{t+\tau} - f_{t+\tau}$). This corresponds to a t -type test of $\alpha = 0$, conditional on the assumption that $\beta = 1(\alpha = 0|\beta = 1)$. To provide an intuition for the two alternative tests, a constant forecast that coincides with the unconditional outcome mean (μ_x) and therefore exhibits a zero mean forecast error would be considered well calibrated, or rather *unbiased*, according to the direct test of a zero mean of the probability forecast error ($x_{t+\tau} - f_{t+\tau}$). However, one would reject the null hypothesis $\alpha = 0$ and $\beta = 1$, as the constant forecast exhibits no correlation with the outcome variable. To avoid confusion of terminology, we refer to the sufficient condition ($\alpha = 0$ and $\beta = 1$) as perfect calibration and to the necessary condition ($\alpha = 0|\beta = 1$) as a test of unbiasedness.

While the hypothesis of a perfectly calibrated forecast implies some positive correlation between the outcome's occurrence and the associated probability forecast, it is also of interest to independently test directly for such a correlation in its own right. Galbraith and van Norden (2012) point out that zero resolution, as defined in (1), requires $\beta = 0$. Under this restriction, the outcome's occurrence and its probability forecast are uncorrelated and, hence, the forecasts do not provide any signal on the event's likelihood of occurrence, i.e., $\mu_{x|f} = \mu_x$. Importantly, this test of zero resolution does not distinguish whether the outcome and the forecast probabilities are positively or negatively correlated. Clearly, however, it is of interest to examine the sign of the estimated β with good forecasts also exhibiting a positive correlation with the outcome. Hence, we focus on a simple one-sided t -test against the null hypothesis of "zero or negative signaling power" or $\beta \leq 0$. Rejection of this null suggests that the forecasts are useful, as they combine the idea of positive resolution with a non-negative correlation between the forecast and the event's occurrence. This latter one-sided test has the additional advantage that in small samples it is necessarily more powerful compared with the two-sided test of zero resolution ($\beta = 0$) studied in Galbraith and van Norden (2012).

Ideally, for good probability forecasts, we would want to *accept* both the tests of perfect calibration but *reject* the restrictions implied by zero or negative signaling power. In testing these hypotheses with the aggregate data, we estimate (2) by OLS whilst controlling for the serial correlation induced by the multi-period nature of

the forecast by using the correction to the standard errors of the parameters suggested by Newey and West (1987). However, Pesaran and Timmermann (2009) have recently highlighted that tests of the significance of ρ_{fx} (the correlation between x and f) are potentially distorted due to the clustering and serial correlation in the observed values of x . They propose a corrected test based on a dynamically augmented reduced rank regression which Lahiri and Wang (2013) have recently applied in examining the resolution of U.S. SPF forecasts. Hence, as an additional robustness check, we supplement the basic version of the one-sided test with the correction suggested by Pesaran and Timmerman (2009). The correction is implemented by a regression of $(\rho_{fx}/\sigma_x^2)x$ on f , augmented by the lagged values of the dependent and the independent variable.⁵ As these two additional regressors contain information about the lagged forecast error, the Pesaran and Timmermann (2009) specification tests for the usefulness of the probability forecast conditional on such lagged information.

2.2 Panel Analysis

In extending the analysis of equation (2) to individual-level data, it is necessary to consider a more general model. In particular, heterogeneity in individual forecaster skill could be associated with different intercept and slope parameters across forecasters. In addition, as highlighted in earlier studies of surveyed point forecasts by Davies and Lahiri (1995), Clements, Joutz, and Stekler (2007), and earlier in Keane and Runkle (1990), an additional complication arises with individual-level data due to the role of aggregate macroeconomic shocks which can result in strong co-movement in forecast errors across individuals. A failure to account for such correlation could bias the estimated coefficients and standard errors and hence render unreliable any inference on the attributes of the SPF probability forecasts. In controlling for these aspects, we employ a GMM procedure similar to that described in Davies and Lahiri (1995), which generalized the approach of Keane and Runkle (1990) to allow for individual forecaster heterogeneity. In particular, we estimate a benchmark panel model that is equivalent to (2) and pools the

⁵See footnote 11 in Lahiri and Wang (2013).

estimated coefficients across individuals but allows for a very flexible structure of the residuals' variance-covariance matrix. The panel version is given in (3) below.

$$x_{t+\tau} = \alpha_+ \beta f_{i,t+\tau} + \varepsilon_{i,t+\tau} \quad (3)$$

Using X to denote the $NT \times 1$ matrix containing N stacked output variables, and F to denote the corresponding $NT \times 2$ matrix for the regression constant and the stacked individual probability forecasts, the vector of estimated regression parameters ($\hat{\beta}_{GMM} = [\alpha, \beta]$) and their variance-covariance matrix (V) are given by $\hat{\beta}_{GMM} = \left[F'Z(Z'\hat{\Omega}^{-1}Z)^{-1}Z'F \right]^{-1} \left[F'Z(Z'\hat{\Omega}^{-1}Z)^{-1}Z'X \right]$ and $Var[\hat{\beta}_{GMM}] = (F'Z)^{-1}(F'\hat{\Omega}^{-1}F)(Z'F)^{-1} = V$, where Z is the set of instruments (see Keane and Runkle 1990). In our implementation, we take $Z = F$, and the estimators are given by $\hat{\beta}_{GMM} = \left[F'F(F'\hat{\Omega}^{-1}F)^{-1}F'F \right]^{-1} \left[F'F(F'\hat{\Omega}^{-1}F)^{-1}F'X \right]$ and $Var[\hat{\beta}_{GMM}] = (F'F)^{-1}(F'\hat{\Omega}^{-1}F)(F'F)^{-1} = V$.

In deriving the above GMM estimators, an essential step is the derivation of a correct residual covariance matrix ($\hat{\Omega}$). In specifying the elements of $\hat{\Omega}$, we follow Davies and Lahiri (1995) and allow for the variance of the regression residuals to potentially differ across individuals, i.e., $Var(\hat{\varepsilon}_{i,t}) = \sigma_i$ for $i = 1, \dots, N$. This approach generalizes the earlier approach in Keane and Runkle (1990) which had assumed homogeneity of error variances across individuals. Within our pooled panel regression framework where the estimated slope and intercept parameters are assumed constant across individuals, such a homogeneity assumption is unattractive, as it implies that no forecaster is systematically better than any other forecaster. Although we have no prior information that would lead us to believe some forecasters are systematically better than others, this later assumption is better tested rather than assumed a priori. Appendix 1 outlines precisely the structure of the assumed variance-covariance matrix. Moreover, as described below in section 2.3, in reporting our results we provide an explicit test of the homogeneity of the estimated coefficients and error variances across individuals following procedures suggested by Hsiao (2003) and Davies and Lahiri (1995), respectively.

2.3 *Specification Tests and Forecast Diagnostics in the Panel Regression*

In this subsection, we describe some of the forecast diagnostic tests that we can apply to our panel regressions. A first important test is of our maintained hypothesis that the intercept and slope parameters of equation (3) do not vary across individuals. We perform this test of our basic model following Hsiao (2003). Allowing for heterogeneity in the forecast bias (α_i) and the slope parameter (β_i), we commence with an unrestricted representation of the model which is given by (4).

$$x_{t+\tau} = \alpha_i + \beta_i f_{i,t+\tau} + \varepsilon_{i,t+\tau} \quad (4)$$

In (4) each of the coefficients are estimated separately by OLS. A test of the validity of the pooled model, proposed by Hsiao (2003), is based on the difference of the sample sum of squared residuals (RSS) of the unrestricted model and the pooled model ($\alpha_i = \alpha$ and $\beta_i = \beta$ for $i = 1, \dots, N$). In particular, the test statistic (F) given in (5) below has a Fisher distribution with $2(N-1)$ and $(NT - 2N)$ degrees of freedom under the null hypothesis of homogenous intercepts and slopes (i.e., $\alpha_i = \alpha$ and $\beta_i = \beta$ for $i = 1, \dots, N$).

$$F = (RSS^{\text{pooled}} - RSS^{\text{unrestricted}}) \times \frac{(RSS^{\text{pooled}} - RSS^{\text{unrestricted}})/2(N-1)}{RSS^{\text{unrestricted}}/(NT-2N)} \quad (5)$$

Rejecting the null hypothesis implies that the pooling assumption is not valid because heterogeneous parameters are required to explain the variability in the panel data set. Hence, this test provides direct empirical evidence on the validity of the pooled coefficients implicit in our benchmark model (3). In reporting our results below, we also explore directly the estimated heterogeneity in individual parameter estimates when relevant.

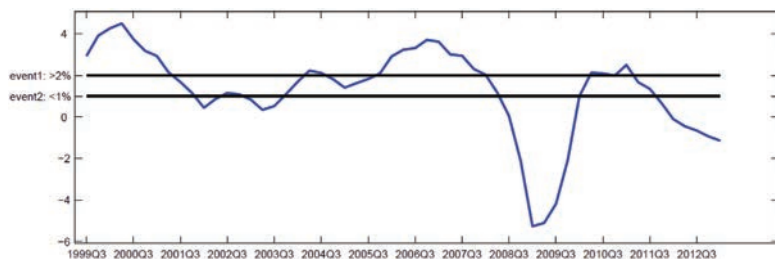
Conditional on the pooled model's assumptions, it is also possible to examine tests of perfect calibration and positive signaling power discussed previously. As in the aggregate-level analysis, the test of perfect calibration implies a joint restriction on the pooled model's two parameters under the null hypothesis and can be implemented as a Wald test using the χ^2 distribution with two degrees of freedom. The alternative, less restrictive test for unbiasedness is

implemented as a t -test on the constant parameter (α) in the regression $x_{t+\tau} - f_{i,t+\tau} = \alpha_i + \varepsilon_{i,t+\tau}$. In the case of the test against the null hypothesis of zero or negative signaling power ($\beta \leq 0$), we can use a simple one-sided t -type test with $N^*T - 2$ degrees of freedom. In addition, given that our general model allows for individual-specific terms in the $\hat{\Omega}$ matrix, we can also test for the homogeneity of the error variances across individuals. Following Davies and Lahiri (1995), such a test can be performed by regressing $\hat{\varepsilon}_{i,t}^2$ on a constant and $N - 1$ individual dummies. The resulting R^2 multiplied by NT is distributed χ_{N-1}^2 under the null hypothesis of homoskedasticity with respect to the cross-sectional dimension, i.e., $\text{Var}(\hat{\varepsilon}_{i,t}) = \sigma_i = \sigma$ for all $i = 1, \dots, N$.

3. Data: Event Forecasts from the SPF

In this section we provide a descriptive review of the event forecasts we evaluate using the methods described in section 2. The complete underlying micro data set can be downloaded at <http://www.ecb.europa.eu/stats/prices/indic/forecast/html/index.en.html>. Our analysis is based on the one- and two-year horizon density forecasts for euro-area real output growth and consumer price inflation, with these variables being measured, respectively, by euro-area gross domestic product (GDP) and the Harmonised Indicator of Consumer Prices (HICP) and published by Eurostat, the statistical agency of the European Union. The analysis is conducted using a filtered SPF data set which excludes *irregular* respondents as described in Genre et al. (2013) and draws on the quarterly rounds of the SPF conducted over the period 1999:Q1–2013:Q2. While such a fifteen-year period provides sufficient data to conduct our empirical tests, it can nonetheless only provide some first tentative signals on the quality of information that can be extracted from the survey. In particular, our results are subject to the significant caveat that they may be sensitive to the arrival of more data—a point emphasized in previous evaluations of the U.S. SPF conducted by Diebold, Tay, and Wallis (1999) and Croushore (2010).⁶ As such, the data comprises

⁶Croushore (2010) finds in particular that the apparent bias found in earlier studies of the inflation forecasts in the U.S. SPF has dissipated as more data has become available. Similarly, the study by Diebold, Tay, and Wallis of the

Figure 1. Outcome for Target Variable: GDP Growth

a cross-sectional dimension of twenty-four to twenty-six individual forecasters, depending on the particular forecast variable or horizon. This set of individual respondents represents a subset of regularly responding forecasters based on a filtering rule that excludes those forecasters who have missed more than four consecutive survey rounds. As a result, the data set is an unbalanced panel with the precise number of time-series observations varying at the individual level, depending on how often a given individual has not submitted a response to the survey. As discussed in Genre et al. (2013), our focus on regular participants in the survey is motivated by the need to avoid sampling bias whereby individual forecaster performance might be excessively driven by a failure to reply during a particular business-cycle period. Our focus on regular respondents also ensures that the number of time-series observations for a given individual is maintained at a reasonably high level (an aspect that is important given the small-sample caveat highlighted above). More complete descriptions of the SPF data set, including a description of its panel dimension, are given in Bowles et al. (2010) and Genre et al. (2013). Garcia (2003) provides an earlier “bird’s-eye” description of the ECB SPF.

Figures 1 and 2 provide time-series plots of the outcomes for the two forecast target variables, where horizontal lines are used to identify the lower and upper fixed thresholds of 1 percent and 2 percent,

U.S. SPF pointed to time variation in density forecast calibration error. In line with this, it cannot be excluded that some of our findings could be reversed by future researchers who can exploit sample sizes greater than the fifteen years that are currently available.

Figure 2. Outcome for Target Variable: HICP Inflation

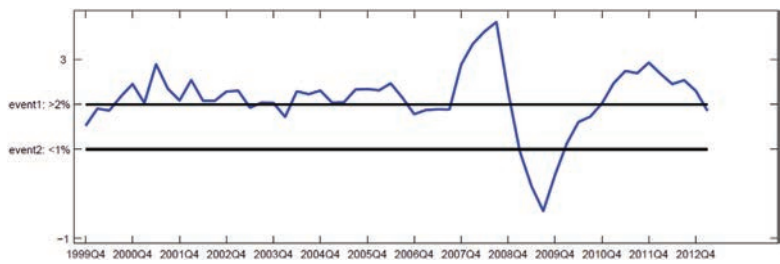


Table 1. GDP Growth Events and Probabilities: Summary Statistics

Event	T	Σx_t	Σx_t^*	μ_x	μ_f	$\rho_{x,f}$
$H = 1$						
GDP Growth > 2%	55	24	4	0.44	0.41	0.43
GDP Growth < 1%	55	18	4	0.33	0.22	0.64
Higher GDP Growth	55	23	5	0.42	0.48	0.66
$H = 2$						
GDP Growth > 2%	51	20	4	0.39	0.58	-0.15
GDP Growth < 1%	51	18	4	0.35	0.09	-0.07
Higher GDP Growth	51	26	3	0.51	0.59	0.74

respectively. To complement this graphical information, tables 1 and 2 provide average sample period information on the events and the forecast probabilities. In the case of GDP growth, for example, the chart indicates four occasions during which growth exceeded the 2 percent threshold we use for the analysis. This is confirmed also in table 1, where for this event we observe that x_t has taken a value of unity in twenty-four of the total of fifty-five outcomes included in our sample (i.e., $\Sigma x_t = 24$).⁷ However, as discussed earlier, x_t

⁷While the SPF probabilities are fully real time in nature, our analysis conditions on the 2013:Q3 vintage for the HICP and GDP outcome series in defining our event outcome series x_t . An additional element of uncertainty for the evaluation relates to the possible impact of alternative data vintages for the occurrence and non-occurrence of the outcome. This may be most relevant for the GDP

**Table 2. Inflation Events and Probabilities:
Summary Statistics**

Event	T	Σx_t	Σx_t^*	μ_x	μ_f	$\rho_{x,f}$
$H = 1$						
Inflation > 2%	54	37	6	0.69	0.36	−0.06
Inflation < 1%	54	4	1	0.07	0.09	−0.03
Higher Inflation	54	31	7	0.57	0.43	0.59
$H = 2$						
Inflation > 2%	50	36	5	0.72	0.39	−0.51
Inflation < 1%	50	4	1	0.08	0.07	−0.30
Higher Inflation	50	29	6	0.58	0.47	0.62

exhibits high persistence and switches from zero to unity on only four occasions, $\sum x_t^* = 4$ where $x_t^* = 1$ if $x_t = 1$ and $x_{t-1} = 0$, and 0 otherwise. Similarly, there were also four occasions when growth fell below the lower threshold of 1 percent, most notably during the Great Recession of 2008 and 2009. At the one-year horizon, the summary statistics in table 1 illustrate that there were five occasions where the GDP growth outcome that emerged was higher than the current growth rate observed at the time the survey was carried out.⁸

In the case of inflation, the pattern is somewhat different, with annual inflation being quite often above the 2 percent threshold we

events, while inflation in the euro area has been less subject to revision (see, for example, the recent study by Kenny, Kostka, and Masera 2014a). We have therefore computed many of the tests we report later in the paper using the first estimate to define the event outcome, and overall we have reached broadly similar conclusions to those that are reached based on the 2013:Q3 vintage.

⁸In the empirical analysis even for the fixed threshold events, the outcome’s frequency of occurrence can change depending on the forecast horizon because the sample size changes. For the two-year-ahead forecast horizon, the sample size for the GDP forecasts is reduced to fifty-one compared with fifty-five for the one-year-ahead forecasts (for inflation the corresponding reduction is from fifty-four to fifty). In the case of the direction-of-change forecasts, the statistics on the outcome’s occurrence/non-occurrence necessarily change with the horizon, which is a defining characteristic of the event. In other words, higher GDP growth/inflation in two years’ time is a conceptually distinct event from higher GDP growth/inflation in one year’s time.

use for this study. For example, as shown in table 2, in thirty-seven of the fifty-four periods studied, inflation was above 2 percent, i.e., 69 percent of the time when measured over the one-year horizon (i.e., $\mu_x = 0.69$). Given this outcome, if the probability assessments of SPF participants are well calibrated, we might expect to observe relatively high probabilities for this event. As shown in table 2, the average probability assigned by SPF respondents stood at 36 percent over the sample ($\mu_f = 0.36$), providing some first-pass evidence that they may have under-estimated this probability. In contrast, the below 1 percent outcome for inflation has occurred only once during the sample (during the 2009 downturn linked to the recent financial crisis), although this occurrence persisted for four quarters. For aggregate analysis, there are considerable limitations and caveats that apply when empirically testing the information content of expert probability assessments, given that we observe the event only once in the sample. In particular, such tests are likely to have relatively modest power and may be biased. However, given that we employ panel data capturing imperfectly correlated expert assessments, in the subsequent analysis we are able to conduct statistical inference concerning the ability of macroeconomists to assess the likelihood of even these relatively infrequently occurring events. Finally, reflecting also the tendency to observe higher inflation outcomes more frequently than lower ones, the direction-of-change event that inflation turns out to be higher than the level observed at the time of the survey has also occurred quite frequently.

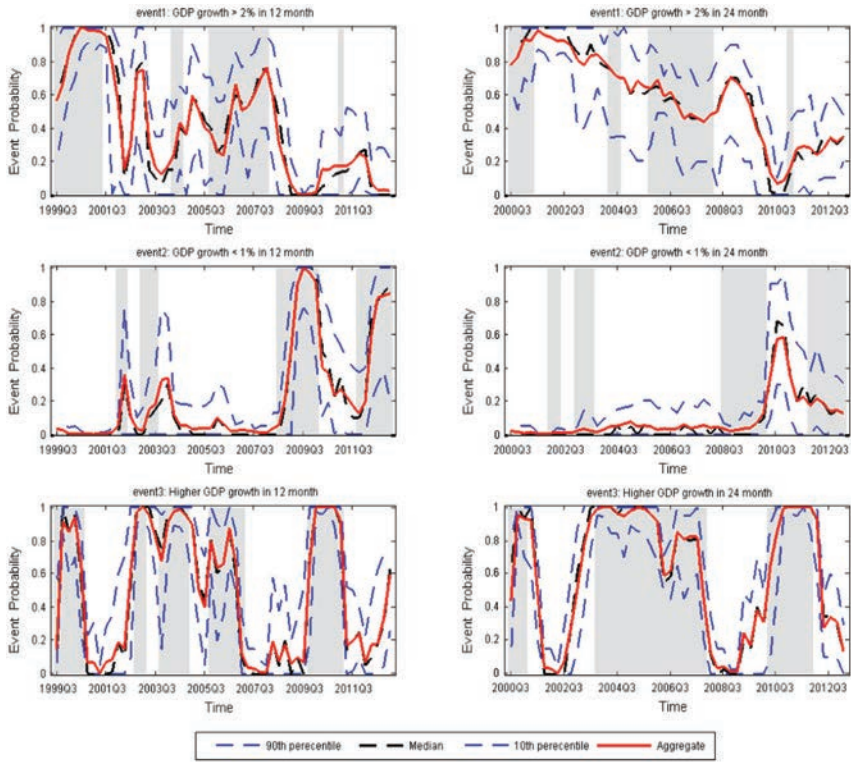
To conduct our analysis, we also need to extract the cumulative probabilities for the three events that we analyze. SPF respondents submit their replies in the form of discrete histograms assigning probabilities to a set of intervals representing possible outcome ranges for the target variable. In addition, at the extremities of these histograms, the assigned probabilities relate to open intervals. To extract the event probabilities from the SPF data, we make the assumption that the probabilities within a given range are uniformly distributed within that range. Without further information on the possible distributional perceptions of survey respondents, the assumption of uniformity seems the most reasonable. An alternative approach would be to fit specific continuous densities to the individual-level data and derive associated event probabilities from

them. However, such an approach could involve the introduction of substantial measurement error. In addition, regarding the open intervals at the edges of the histogram, these are assumed to be closed intervals of equal width to the other surveyed intervals. We have conducted some sensitivity analysis to an alternative assumption that the open intervals are twice the width of the closed intervals and found no notable impact. This is related to the properties of the survey data where in fact, at the individual level, it is often the case that either a very small or a zero probability is assigned to these open intervals.

Figures 3 and 4 present the probability forecasts for the three different events for GDP growth and inflation at both forecast horizons. The probabilities are depicted showing the median probability together with the 10th and 90th percentiles extracted from the cross-section of individual surveyed densities. Also reported is the probability extracted from the aggregate SPF density, which in general is often very close to the median probability. For each chart, we also depict (using shading) the periods during which the event in question actually occurred. From the charts, the direction-of-change assessments appear to correlate quite well with the actual occurrence of the events. Indeed, as also shown in tables 1 and 2, all of the direction-of-change forecasts exhibit a positive correlation with the outcome, i.e., ρ_{xf} is 0.66 and 0.74 for GDP $H = 1$ and $H = 2$, respectively, and 0.59 and 0.62 for the corresponding inflation forecasts. The positive correlation suggests that even at the longer horizons experts can accurately gauge the cyclical tendency of the economy compared with its current state. At the same time, these simple correlations need to be interpreted with due caution and warrant a more robust econometric evaluation, which we provide in the next section.

This first visual inspection of the probability forecasts in figures 3 and 4 suggests a less clear correspondence between the expert probability assessments for the high/low fixed threshold events. A good example of this is the probability assessment for low inflation. In the case of the one-year horizon, the probabilities for this event (figure 4) appear to be lagging, starting to rise only after the event in question had actually occurred. In the case of the two-year assessment, this lagging pattern is even more evident, with the probabilities only starting to rise after the low inflation outcome had

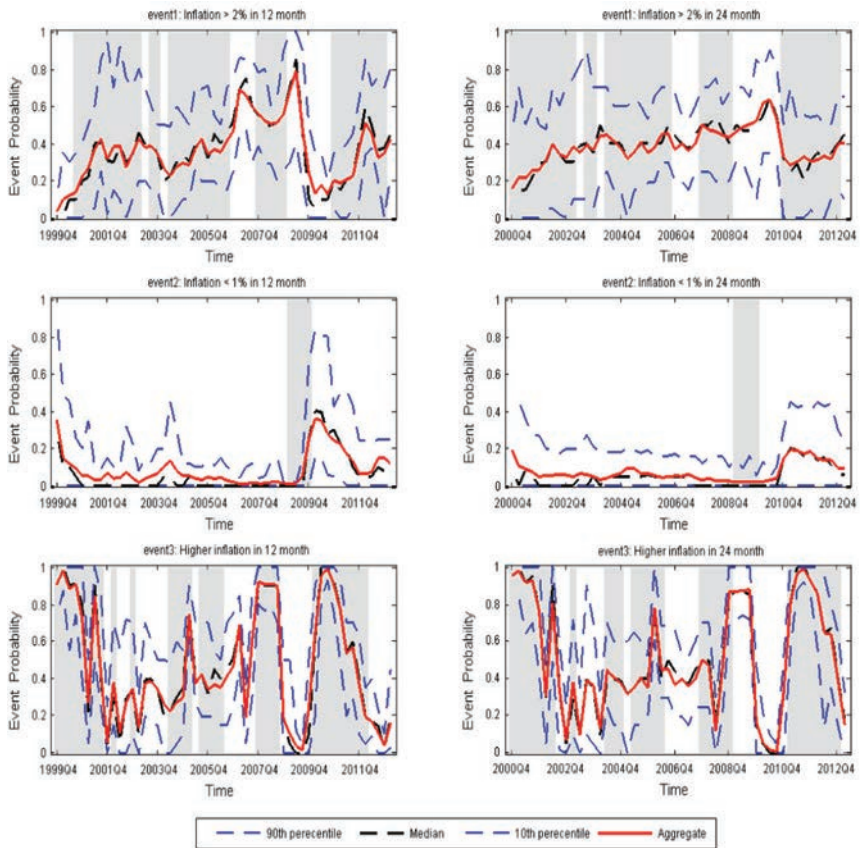
Figure 3. Probability Forecasts (Median, 10th, and 90th Percentiles) for GDP Events



Notes: The shaded region indicates the periods in which the event related to each corresponding probability forecast actually occurred.

completely passed. This graphical evidence of the weaker relationship for the fixed threshold events is also highlighted by a relatively small negative correlation between the aggregate probability forecasts and the observed outcome in table 2. The graphical evidence for the fixed threshold GDP events also suggests relatively limited and/or lagging signaling power of the SPF probability assessments. An exception is perhaps the one-year-ahead GDP predictions for both high and low outcomes which, as observed in table 1, also exhibit a positive correlation of 0.43 and 0.64 with the high and low outcomes, respectively.

Figure 4. Probability Forecasts (Median, 10th, and 90th Percentiles) for Inflation Events



Notes: The shaded region indicates the periods in which the event related to each corresponding probability forecast actually occurred.

4. Evaluation Results for SPF Event Forecasts

In this section, we exploit the QPS decomposition and calibration and resolution tests discussed previously in order to shed more robust econometric evidence on the information content of these attributes of the SPF densities. We first report the QPS decomposition, providing evidence of miscalibration and signaling power.

We then report more formal tests of perfect calibration and unbiasedness and tests of zero or negative signaling power using both aggregate and pooled individual-level data. Finally, we explore the heterogeneity in the SPF panel in more detail. We consider first the results for the GDP events and then subsequently turn to inflation.⁹

4.1 GDP Growth Events

Table 3 reports the QPS decomposition for each of the three GDP events. The decomposition is based on the aggregated probabilities as published on a quarterly frequency by the ECB. The QPS statistics indicate that the SPF densities perform less well at capturing the more extreme threshold events, whilst the direction-of-change predictions have lower QPS and hence performed better. For all three events at short horizons ($H = 1$), the aggregate SPF probabilities appear close to perfectly calibrated (as indicated by a very small calibration error). They also exhibit some resolution, which is particularly strong for the direction-of-change forecasts. In contrast, the signaling information for the upper and lower fixed threshold events is smaller. At the two-year horizon, there is a notable increase in the miscalibration for the fixed threshold event forecasts, while the calibration error for the direction-of-change forecast continues to be very small. The latter forecasts also continue to possess useful signaling information, as reflected in the estimated resolution even at the longer horizon. Overall, therefore, the QPS decomposition suggests the SPF densities for GDP are most informative at short horizons and provide less reliable information about specific high/low GDP growth events that lie further ahead in the future. In contrast, the direction-of-change information appears informative, even at longer horizons. We can, however, provide more formal evidence on this using the regression-based tests of perfect calibration and signaling power described in section 2.

Table 4 reports the results from the estimation of equation (2) based on the aggregate probabilities. The estimation results tend to confirm the observations made above. In particular, at short horizons we are unable to reject the hypothesis of perfect calibration

⁹Appendix 2 presents some sensitivity analysis where we vary the fixed thresholds used to define the high and low outcome events.

Table 3. Decomposition of Quadratic Probability Score: GDP Events

Event	QPS	=	Uncertainty	+	Calibration Error	–	Resolution
	$E[X-f]^2$		σ_x^2		$E_f[\mu_{x f}-f]^2$		$E_f[\mu_{x f}-\mu_x]^2$
$H = 1$							
GDP Growth > 2%	0.38		0.49		0.01		–0.12
GDP Growth < 1%	0.25		0.37		0.01		–0.13
Higher GDP Growth	0.24		0.50		0.00		–0.27
$H = 2$							
GDP Growth > 2%	0.71		0.47		0.24		0.00
GDP Growth < 1%	0.54		0.39		0.17		–0.02
Higher GDP Growth	0.20		0.49		0.00		–0.29

Table 4. Tests of GDP Growth Events: Aggregate Densities

				(1) $H_0:$	(2) $H_0:$	(3) $H_0:$	(4) $H_0:$
	T	α	β	$\alpha = 0, \beta = 1$	$\beta \leq 0$	$\alpha = 0 \beta = 1$	$\gamma \leq 0$
<i>Forecast Horizon ($H = 1$)</i>							
GDP Growth > 2%	55	0.15 (0.12)	0.70 (0.22)	0.327	0.001	0.755	0.034
GDP Growth < 1%	55	0.11 (0.07)	1.02 (0.10)	0.244	0.000	0.106	0.001
Higher GDP Growth	55	0.01 (0.07)	0.85 (0.14)	0.542	0.000	0.425	0.000
<i>Forecast Horizon ($H = 2$)</i>							
GDP Growth > 2%	51	0.56 (0.25)	-0.29 (0.37)	0.000	0.787	0.212	0.013
GDP Growth < 1%	51	0.38 (0.15)	-0.25 (0.61)	0.038	0.663	0.016	0.095
Higher GDP Growth	51	-0.06 (0.06)	0.97 (0.11)	0.521	0.000	0.306	0.002

Notes: Estimates of α and β are obtained from the OLS regression of $x_{t+\tau} = \alpha + \beta^*f_{t+\tau} + \varepsilon_{t+\tau}$, where $f_{t+\tau}$ denotes the probability forecasts extracted from the equal weighted aggregate SPF density and $x_{t+\tau}$ the corresponding binary outcome indicator. Newey-West standard errors correcting for serial correlation are given in parentheses. Columns 1 and 2 report p-values for the indicated hypotheses. The p-values in column 3 are obtained from the OLS regression $x_{t+\tau} - f_{t+\tau} = \alpha + \varepsilon_{t+\tau}$. The p-values in column 4 are obtained from the OLS regression $\theta x_{t+\tau} = c(1 - \phi) + \gamma^*f_{t+\tau} - \phi\gamma^*f_{t+\tau-1} + \phi\theta^*x_{t+\tau-1} + \varepsilon_{t+\tau}$, where $\theta = \rho_{x,f}/\sigma_x^2$.

($H_0: \alpha = 0, \beta = 1$) for all three event forecasts. Indeed, at this horizon, the parameter estimates for the low growth outcomes and the direction-of-change forecast are remarkably close to the values under the null hypothesis. Not surprisingly, therefore, the Wald test is unable to reject. For the three events considered, we can also reject the null hypothesis of negative or zero signaling power ($H_0: \beta \leq 0$) at $H = 1$. This tends to confirm the useful signaling information in the SPF densities for GDP at short horizons, both for the direction-of-change and the fixed threshold events. However, at longer horizons, the above findings are reversed. For both high and low fixed threshold events, we reject the null of perfect calibration at the 5 percent level. The null of zero or negative signaling power cannot be rejected for both the relatively high and low growth thresholds at the two-year horizon. The estimates of β for these events at this longer horizon also tend to be negative, implying that the probability forecasts tend to fall when these high/low GDP growth events occur and vice versa. Such an inverse correlation points to the relatively poor information content of the aggregate densities for such events at longer horizons. For the direction-of-change forecast at this longer horizon, however, we continue to accept the null of perfect calibration and reject the hypothesis of zero or negative signaling power. Hence, even at this longer horizon, econometric results tend to confirm some important information content of the SPF densities for the direction of change in GDP.

In general, the aggregate analysis discussed above points to a high degree of information content in SPF probability forecasts for the direction of change in GDP growth at both short and longer horizons. This finding also appears robust when one considers other variants of the tests discussed in section 2. For example, the direct test of the zero mean in the probability forecast error cannot be rejected, while a test based on the Pesaran and Timmermann (2009) correction for clustering in the outcome variable also rejects the null of zero or negative signaling power in the direction-of-change forecasts. Also, when examining the results of these other tests (reported in columns 3 and 4 in table 4), the evidence highlights the relatively good performance of the fixed threshold event forecasts at short horizons ($H = 1$). Indeed, in contrast to our baseline result, the Pesaran and Timmermann (2009) specification suggests that the hypothesis of zero or negative signaling power for the high/low GDP growth

event forecasts is rejected even at longer horizons. This result implies that the probability forecast contains some information beyond the information that is contained in the lagged outcome and the lagged forecast.¹⁰

Table 5 reports the equivalent regression results for all three GDP event forecasts but based on the unbalanced panel of individual responses and including a correction for both serial correlation and aggregate shocks (using the GMM estimation procedure described in section 2 and appendix 1). In columns 1 and 2 we report the results of the tests for the commonality of coefficients in the panel regression model and the homogeneity of error variances, respectively. In general, these tests support the validity of pooling coefficients for all events considered and at both horizons. Similarly, controlling for aggregate shocks, we are unable to reject the assumption that the error variances are equal across forecasters. We interpret these findings as supporting the idea that the heterogeneity in SPF forecasts for GDP cannot be seen as evidence of significant differences in forecast performance of individual forecasts when viewed from an ex post perspective. This result is somewhat in line with the findings in a recent study by D'Agostino, McQuinn, and Whelan (2012) in relation to the point forecasts in the U.S. SPF. In particular, they find limited evidence for the idea that the best forecasters are actually significantly better than others, though there is evidence that a relatively small group of forecasters perform very poorly. Our finding above is also in line with recent analysis of ECB SPF densities conducted in Kenny, Kostka, and Masera (2014b), who show that for GDP growth a significant improvement in aggregate density performance cannot be achieved by simply excluding those forecasters with a relatively poor track record.

Regarding the key attributes of the forecasts, the panel results in table 5 tend to confirm many of the findings observed in table 4 using the aggregate-level data. For short horizons, the hypothesis that the probability forecasts exhibit zero or negative signaling power is strongly rejected (column 4). At longer horizons, however,

¹⁰Given the small sample available for estimation with the aggregate results and the additional parameters that need to be estimated in the implementation of this test, we would think that some caution is warranted in interpreting this result.

the probability forecast for relatively high and low growth outcomes exhibits no useful signaling power and is inversely correlated with the outcome, as highlighted by negatively estimated slope parameters for these events. Although the low growth event does not correspond to a recession in a classical sense, this latter result for long horizons is broadly in line with Harding and Pagan (2010), who review the literature and empirical evidence on the predictability of recessions and conclude that it is very difficult to predict these events *ex ante*. One interesting feature of the panel results, which compliments the earlier findings based on aggregate probabilities, is the inability to reject the perfect-calibration hypothesis at the individual level for fixed threshold events at the one-year forecast horizon but the strong tendency to reject perfect calibration at the longer two-year horizon (column 3). As was shown with the aggregate-level analysis, the direction-of-change forecasts for GDP growth appear well calibrated at both short and longer horizons, also when basing the analysis on the less demanding hypothesis of a zero mean in the forecast error (column 5). In contrast, for the fixed threshold events, our empirical results point to a quite dramatic deterioration in the information value of density forecasts for real output growth when the forecast horizon is extended. The probability assessments of experts included in the SPF panel thus provide insightful information on the general cyclical evolution of the economy, but at longer horizons they are less informative about the occurrence or non-occurrence of specific high or low GDP growth events.

In assessing the above findings on the information content of the probability forecasts extracted from the ECB SPF, it is interesting to compare them with previous studies using U.S. data. For example, in recent studies of the U.S. SPF, Lahiri and Wang (2006, 2013) study the usefulness of probability forecasts for GDP declines. They find that the shorter-horizon forecasts for GDP growth are more informative than longer-horizon ones—a finding which directly echoes our results for the euro area. At the same time, our results point to noticeable limitations for the probability forecasts for specific fixed threshold events at longer horizons. Lahiri and Wang (2006) highlight in particular the limitations of the probability forecasts for less frequently occurring events such as recessions and point to the role of behavioral factors that may lead forecasters to inappropriately anchor their expectations in difficult forecasting environments.

Similar limitations on the information content of the SPF density forecasts for U.S. output growth are also highlighted in Clements (2010), who finds that not all forecasters update the information in their histograms as new information arrives. Finally, a key aspect of our analysis has been the exploration of heterogeneity in the performance of the densities across individuals. For the GDP densities, our panel results point to the poolability of the information in the density forecasts and hence a relatively low degree of heterogeneity across forecasters' performance when viewed from an ex post evaluation perspective.¹¹ Other studies have found stronger evidence of significant forecaster heterogeneity. For example, studies such as Boero, Smith, and Wallis (2008), Engleberg, Manski, and Williams (2009), and previously Davies and Lahiri (1995) emphasized the significance of such heterogeneity in relation to the individual point forecasts.

4.2 Inflation Events

Table 6 reports the QPS and its associated decomposition for each of the three inflation events at both one- and two-year horizons. The results compare somewhat less favorably with the previous findings for GDP. In particular, the inflation scores are, as a rule, noticeably higher than the corresponding GDP scores. Moreover, SPF probability forecasts for both high and low fixed threshold inflation events show signs of miscalibration—even at the shorter horizons. Similarly, the probability forecasts for the specific fixed threshold events exhibit quite low signaling power, as reflected in low resolution. However, once again, the direction-of-change forecasts appear better calibrated and exhibit higher resolution. This is mainly at the one-year horizon, however, and the evidence that the direction-of-change forecasts for inflation provide useful signals (i.e., high resolution) at longer horizons is weaker.

Tables 7 and 8 report the econometric tests for perfect calibration and zero resolution for each of the three inflation events for the aggregate- and panel-level analysis, respectively. At the aggregate level (table 7), the results for inflation contrast with some of the

¹¹We explore further the empirical evidence on forecaster heterogeneity in section 4.3 below.

Table 6. Decomposition of Quadratic Probability Score: Inflation Events

Event	QPS	=	Uncertainty	+	Calibration Error	-	Resolution
	$E[X-f]^2$		σ_x^2		$E_f[\mu_{x f}-f]^2$		$E_f[\mu_{x f}-\mu_x]^2$
$H = 1$							
Inflation > 2%	0.72		0.44		0.27		0.00
Inflation < 1%	0.54		0.15		0.39		-0.01
Higher Inflation	0.32		0.48		0.02		-0.18
$H = 2$							
Inflation > 2%	0.73		0.42		0.42		-0.11
Inflation < 1%	0.44		0.17		0.29		-0.02
Higher Inflation	0.41		0.50		0.02		-0.10

Table 7. Tests of Inflation Events: Aggregate Density

			(1) $H_0:$	(2) $H_0:$	(3) $H_0:$	(4) $H_0:$	
	T	α	β	$\alpha = 0, \beta = 1$	$\beta \leq 0$	$\alpha = 0 \beta = 1$	$\gamma \leq 0$
Forecast Horizon ($H = 1$)							
Inflation > 2%	54	0.75 (0.26)	-0.17 (0.63)	0.001	0.607	0.003	0.419
Inflation < 1%	54	0.08 (0.08)	-0.07 (0.30)	0.000	0.599	0.806	0.087
Higher Inflation	54	0.14 (0.09)	1.01 (0.13)	0.109	0.000	0.038	0.000
Forecast Horizon ($H = 2$)							
Inflation > 2%	50	1.67 (0.27)	-2.46 (0.72)	0.000	0.999	0.006	0.112
Inflation < 1%	50	0.20 (0.14)	-1.67 (1.12)	0.000	0.932	0.925	0.389
Higher Inflation	50	0.09 (0.08)	1.06 (0.09)	0.040	0.000	0.041	0.000

Notes: Estimates of α and β are obtained from the OLS regression of $x_{t+\tau} = \alpha + \beta^*f_{t+\tau} + \varepsilon_{t+\tau}$, where $f_{t+\tau}$ denotes the probability forecasts extracted from the equal weighted aggregate SPF density and $x_{t+\tau}$ the corresponding binary outcome indicator. Columns 1 and 2 report p-values for the indicated hypotheses. The p-values in column 3 are obtained from the OLS regression $x_{t+\tau} - f_{t+\tau} = \alpha + \varepsilon_{t+\tau}$. Newey-West standard errors correcting for serial correlation are given in parentheses. The p-values in column 4 are obtained from the OLS regression $\theta x_{t+\tau} = c(1 - \phi) + \gamma^*f_{t+\tau} - \phi\gamma^*f_{t+\tau-1} + \phi\theta^*x_{t+\tau-1} + \varepsilon_{t+\tau}$, where $\theta = \rho_{x, f} / \sigma_x^2$.

previous findings observed for GDP. In particular, in column 1, the test of well-calibrated forecasts tends to be rejected for the fixed threshold events at both horizons.¹² Also, we are unable to reject the null of negative signaling power (column 3) for all fixed threshold events. Moreover, as indicated by the estimated β parameters and the tests of negative signaling power (column 2, table 7), the SPF predictions for the fixed threshold inflation events appear to be either uncorrelated or correlate negatively with the occurrence of these events. Hence, in the case of inflation, our results would strongly suggest for users of the SPF information to exercise considerable caution when extracting information on the likelihood of more specific events toward the upper and lower range of inflation outcomes. Turning to the direction-of-change forecasts for inflation, the aggregate results tend to mirror some of the more positive findings that we observed for GDP. In particular, direction-of-change forecasts appear quite informative, as they are better calibrated and exhibit significant positive resolution. Even though the hypothesis of perfect calibration is often rejected at the 10 percent confidence level, the unrestricted estimates of α and β for the direction-of-change forecasts in table 7 correspond closely with their predicted values under the null.

The above aggregate results on the poor calibration of fixed threshold events for inflation are confirmed by the panel analysis in table 8 where the test of perfect calibration (column 3) is rejected for all four fixed threshold events and at both horizons. This finding for inflation compares less favorably with the equivalent analysis for GDP, where the fixed threshold probabilities were shown to be better calibrated at short horizons ($H = 1$). Moreover, and in line with the aggregate-level analysis, as indicated by the estimated β parameters and the tests of negative signaling power (column 4, table 8), the SPF probabilities for the fixed threshold inflation events appear to be either uncorrelated or correlate negatively with the occurrence of these events. Also mirroring the aggregate-level analysis,

¹²For the fixed threshold inflation events, the direct test of the mean probability error in column 3 of table 8 provides a conflicting signal to the test in column 1 of table 8. This discrepancy reflects the large deviation of the unrestricted estimates of β from its value of unity under the null in the direct test. Given the evidence of noticeable miscalibration in table 6, we would tend to trust more the results of the test in column 1.

direction-of-change forecasts appear quite informative, as they are better calibrated and exhibit significant positive resolution. For both forecast horizons, the test of negative signaling power can be rejected in contrast to the equivalent fixed threshold events. Once again, this finding highlights the ability of SPF panel members to capture the broad cyclical direction of inflation but a much weaker ability to correctly diagnose the likelihood of more specific high and low inflation outcomes. Concerning the heterogeneity of the SPF forecasts for inflation, the test results reported in table 8 suggest that the pooling of individual coefficients and error variances is, for most events, not rejected by the data. One interesting exception is the event for high inflation outcomes above 2 percent where the common slope and intercept in the panel model is rejected (although the common error variance is not). These differences in individual forecaster attributes for this event, which is of some interest from a monetary policy perspective, are discussed in more detail in section 4.3 below.

The above evaluation results for the inflation densities from the euro-area SPF point to some noticeable limitations in particular as regards their information about specific relatively high or low inflation outcomes. In interpreting the result for the fixed event thresholds at shorter horizons ($H = 1$), the relatively poor predictive performance compared with the equivalent GDP events warrants particular note. On the one hand, due to more significant data revisions for GDP, one might have expected the inflation probabilities to perform better than the equivalent GDP probabilities. However, this result mirrors previous findings which considered the point forecast performance from the euro-area SPF. As discussed in Genre et al. (2013), an important factor which may explain this is the substantial role for commodity price shocks in driving euro-area inflation over the sample period. As shown in Bowles et al. (2010), the predictive performance of the SPF for core inflation has tended to be somewhat better. In addition, it should be noted that the effective forecast horizon for inflation at $H = 1$ is longer due to the more significant publication lags for the GDP release.¹³ It is also of interest to compare our

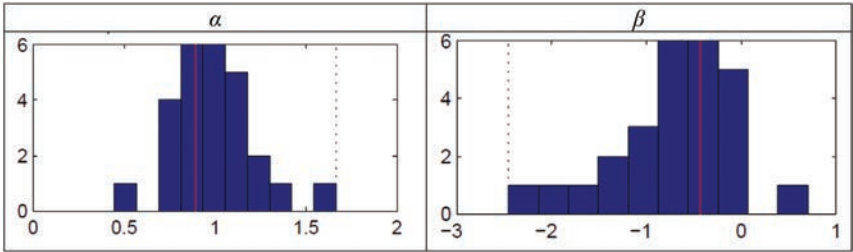
¹³ As the ECB SPF sets the horizon exactly one year after the latest publication release, the forecast target date for GDP is two quarters from the survey quarter, while for inflation it is eleven months.

findings for inflation with some of the existing empirical literature that has analyzed the density forecasts from the SPF for the United States. For example, Clements (2006) used the probability integral transform method and finds that the efficiency of the inflation densities from the U.S. SPF is often rejected. However, echoing our more positive results in relation to the direction-of-change forecasts, he does find that the aggregate inflation densities from the U.S. SPF are conditionally efficient compared with a naïve “no change” prediction. Diebold, Tay and Wallis (1999) have also used the probability integral transform (PIT) to evaluate the inflation densities from the U.S. SPF and conclude that the U.S. SPF densities for inflation are not optimal, as the PITs are non-uniform and serially correlated. Part of the reason for poor density performance may relate to overconfidence in the inflation densities as highlighted for the euro area in Kenny, Kostka and Masera (2014a), a feature also documented by Giordani and Söderlind (2006).

4.3 Exploring Further the Heterogeneity in Individual Event Forecasts

The preceding analysis has focused on documenting the event forecast performance of macroeconomists drawing on tests based on aggregate probability distributions or the individual densities. Such an approach directly addresses the question posed in the title of this paper, as it sheds light on whether or not surveyed densities are informative either when aggregated or at a micro level. For most of the events that we have examined, our analysis has supported the pooling of the individual densities across forecasters in terms of both the estimated slope and the intercept coefficients in the panel regressions. In addition, we are often unable to reject the hypothesis of a common forecast error variance among individual forecasters. An exception was the inflation density forecasts, where the evidence pointed to more statistically significant differences in individual forecast performance, as reflected in heterogeneous slope and intercept coefficients. In this section we explore further the extent and nature of heterogeneity in SPF density forecasts at an individual level. To shed light on this, we report the results from the estimation of equation (3) but allow for variation in the intercept and slope parameters across individuals. As an example, figure 5 provides information

Figure 5. Histogram of Individual-Level Parameter Estimates: Inflation > 2% (H = 2)



Notes: The bars denote the number of forecasters for which the estimated parameter takes the value given on the horizontal axis. The vertical line depicts the estimated parameters based on the panel regressions. The dotted line denotes the estimated parameters based on the probabilities from the aggregate distributions.

on the estimated constant and slope parameter at the individual level for the case of high inflation outcomes above 2 percent at the longer horizon ($H = 2$). It can be recalled from table 8 that for this event the commonality of parameters across individuals was rejected. The figure depicts a histogram measuring on the vertical axis the number of individual forecasters, with the estimated parameter values indicated by the range of values on the horizontal axis. Also reported are the parameter estimates based on the aggregate distributions (indicated by a dotted vertical line and taken directly from table 7) together with the median parameter values (indicated by a solid vertical line). The histograms confirm that the relatively poor calibration of SPF forecasts for this inflation event is broadly shared across the majority of forecasters, as indicated by estimates of α which are consistently above zero. Indeed, all forecasters in the panel have tended to underpredict the occurrence of relatively high inflation outcomes. In terms of signaling power (resolution), the individual-level parameter estimates for β also suggest substantial heterogeneity, with the estimates ranging from below -2 to a small positive value in the case of the forecaster with the highest resolution. Yet, despite the heterogeneity in individual forecast attributes, all forecasters appear to be far away from having well-calibrated probability assessments for this event. A more formal assessment of this is provided below, where we report the share of individual

forecasters for which our different hypotheses on individual forecast attributes can be rejected.

Table 9 summarizes the results of the individual-by-individual regressions for all three events, for both target variables and both forecast horizons. For the hypothesis of perfectly calibrated forecasts, the table reports rejection rates measured as the number of individuals for which the hypothesis is rejected (at the 10 percent level) expressed as a share of the total number of individuals in the panel.¹⁴ Hence, the table provides some summary information on the degree of heterogeneity in density forecast performance. Rejection rates at the upper and lower bounds—i.e., that are close to either 0 percent or 100 percent—indicate a high level of homogeneity in forecast performance, while shares that are away from these bounds highlight possible heterogeneity in forecaster performance. The figures in table 9 confirm that the previous results concerning the relatively well-calibrated predictions for the fixed threshold event of GDP growth < 1 percent at the short horizon ($H = 1$) and the direction-of-change forecasts at both horizons are relatively widely shared across individual forecasters. For example, for the GDP growth < 1 percent event at the short horizon ($H = 1$), we can reject the hypothesis of perfectly calibrated forecasts for only 13 percent of the forecasters in our panel. Similarly, the miscalibration of the fixed threshold event of GDP growth > 2 percent at longer horizons ($H = 2$) is widely shared across many forecasters in the panel, as indicated by rejection rates of 88 percent and 76 percent for the high and low outcomes, respectively. At the same time, some regressions exhibit a somewhat higher degree of heterogeneity. For instance, the hypothesis of perfect calibration of the GDP growth > 2 percent event forecast at the short horizons ($H = 1$) was neither rejected in the aggregate nor in the pooled approach (see tables 4 and 5, respectively). However, the same hypothesis can be rejected for 29 percent of all individual forecasters. Regarding inflation, we also observe that the miscalibration of the above 2 percent event is quite homogenous across all forecasters at both

¹⁴The rejection rates are computed after correcting for the false discovery rate in sequential hypothesis testing as proposed by Benjamini and Hochberg (1995). In this setup, the null is rejected at the $\alpha^{\text{cor}} = 10\% * i/N$ level, i being the individual with the i -th lowest p-value from the set of individual-level regressions.

horizons, as indicated by rejection rates of 88 percent ($H = 1$) and 100 percent ($H = 2$). However, for the inflation outcome below 1 percent and the direction-of-change forecasts, it is interesting to observe that the hypothesis of well-calibrated forecasts is only rejected for a smaller but non-negligible share of forecasters. This points to a stronger degree of heterogeneity in forecaster skill for these specific events.

Table 9 also provides information on the resolution and signaling power of the probability forecasts at the individual level. In particular, it reports the number of individuals for which the one-sided hypothesis $\beta \leq 0$ is rejected (at the 10 percent level), again expressed as a share of the total number of individuals. Also reported are the equivalent rejection rates using the Pesaran and Timmermann (2009) adjustment for the clustering in the binary outcome variable. Rejection of these hypotheses implies that the probability forecasts are informative in the sense that they have non-zero resolution *and* are positively correlated with the event's occurrence. Concerning this important attribute of the SPF probability forecast, we again observe a high degree of homogeneity across individuals with rejection rates that are close to the upper and lower bounds of 100 percent and 0 percent. The previous findings of positive signaling power for the direction-of-change forecasts for both GDP growth and inflation are shown to be very homogenous and shared by most or even all experts in the panel (for both horizons, the rejection rates are above 90 percent for the hypothesis that $\beta \leq 0$). Equally, the lack of any positive signaling power for the fixed threshold inflation events is shared by nearly all forecasters for both horizons (as indicated by rejection rates that are below 10 percent). This picture is somewhat modified by the results of the Pesaran and Timmermann test for which the rejection rates are further away from the 100 percent and 0 percent bounds. For instance, we find lower rejection rates for the short-horizon GDP forecasts compared with the equivalent test for zero or negative resolution ($\beta \leq 0$). These conflicting indications of potential significant heterogeneity in the SPF forecasts may reflect the small sample that is currently available for empirical analysis. For example, in the case of the Pesaran and Timmermann test, four parameters need to be estimated for each individual, with a maximum of fifty-four observations. Nonetheless, the evidence of more significant heterogeneity in individual forecast performance

will warrant further investigation in the future particularly as more data become available.

The above analysis of density forecast heterogeneity demonstrates that when such forecasts are poorly calibrated or non-informative, this poor performance often tends to be widely shared across the sample of forecasters in our panel. Correspondingly, when we observe some information content in the density forecasts—such as is the case for direction-of-change forecasts for both GDP growth and inflation or for high and low GDP growth outcomes but only at short horizons—such forecaster skill is often quite evenly distributed across most individuals. At the same time, the extent of heterogeneity in individual forecaster skill can also depend on the specific forecasting task at hand. For some specific events and horizons, and particularly for the calibration of the inflation forecasts, we observe rejection rates that are suggestive of more noteworthy heterogeneity. This evidence on forecaster heterogeneity strikes a note of caution for the strong poolability that was observed in many of the panel regressions. It also accords more closely with other studies of the U.S. SPF such as Engleberg, Manski, and Williams (2009) and previously Davies and Lahiri (1995) which, although using different methods and perspectives, had emphasized the heterogeneity in individual SPF replies. Future research might therefore consider carefully the extent to which such heterogeneity in forecaster skill can be exploited in order to improve the usefulness of surveys such as the SPF. For example, it may be possible to enhance aggregate density performance by excluding some forecasters whose densities exhibit persistently poor calibration, weak signaling power, and low information content.¹⁵

¹⁵This suggests the need to investigate alternatives to the current practice of taking an equal weighted average when aggregating individual SPF replies. For example, Jore, Mitchell, and Vahey (2010) have identified gains from combined density forecasts that weight more highly the more informative component densities. When the number of density forecasts available is large, estimation of individual density weights can become computationally burdensome. Recently, however, Conflitti, De Mol, and Giannone (2012) have proposed methods to estimate optimal combination weights and have applied their method to the euro-area SPF, finding some modest gains.

5. Conclusions

In this paper we have attempted to shed further light on the information content of the density forecasts of macroeconomists collected in surveys. Our analysis yields a number of findings of relevance to central banks, and others, making use of SPF results to inform their decisions. In general, and bearing a small-sample caveat in mind, we have observed relatively low information content in the SPF density forecasts for relatively high and low outcome events. In contrast, the SPF densities appear considerably more informative concerning more central tendencies in the forecast target variable, as reflected in their probabilistic assessments of its likely future directional change. This result confirms the ability of the survey participants to capture in their density forecasts normal cyclical fluctuations (e.g., mean reversion). The evidence we uncover would thus support the case to monitor direction-of-change indicators from surveys such as the ECB's SPF. In general, our analysis also points to a rather high degree of homogeneity in forecaster skill along some of the dimensions that we have examined and when assessed from an ex post perspective. At the same time, we have uncovered some evidence to suggest that the heterogeneity in forecaster skill can also depend on the specific forecasting task. Future research might therefore consider carefully the extent to which such heterogeneity can be exploited in order to improve the usefulness of surveys such as the SPF.

Appendix 1

In line with the discussion in section 2.2, we construct our variance-covariance matrix under the following simplifying assumptions:

$$E[\varepsilon_{i,t+l}\varepsilon_{j,t+m}] = \left\{ \begin{array}{ll} \sigma_{i|l-m|} \geq 0 & \text{for } i = j \text{ and } |l-m| \leq \tau \\ \delta_{l-m|} \geq 0 & \text{for } i \neq j \text{ and } |l-m| \leq \tau \\ 0 & \text{otherwise} \end{array} \right\}. \quad (6)$$

In (6), both individual-specific error variances ($\sigma_{i,0}$) and the covariances of the lagged errors up to order τ ($\sigma_{i,1}, \dots, \sigma_{i,\tau}$) are allowed to be different across each individual in the SPF panel. In contrast,

$\delta_0, \dots, \delta_\tau$ reflect coincident and lagged covariances across individuals that are due to aggregate shocks and, hence, are common to all individual forecasters. The corrected estimate of $\hat{\Omega}$ thus has several off-diagonal non-zero elements capturing the heterogeneity as well as the time and cross-sectional correlation in the residuals of equation (2). More specifically, it takes the form given in equation (7).

$$\hat{\Omega} = \begin{bmatrix} A_1 & B & \cdot & \cdot & \cdot & B \\ B & A_2 & B & \cdot & & \\ B & B & A_3 & B & \cdot & \\ \cdot & \cdot & \cdot & \cdot & & \\ B & \cdot & \cdot & \cdot & & A_N \end{bmatrix} \quad (7)$$

For instance, in the case of the four-quarter-ahead probability forecast ($\tau = 4$), the matrices A and B are constructed as follows:

$$A_i = \begin{bmatrix} \sigma_{i,0} & \sigma_{i,1} & \sigma_{i,2} & \sigma_{i,3} & \sigma_{i,4} & 0 & \cdots & 0 \\ \sigma_{i,1} & \sigma_{i,0} & \sigma_{i,1} & \sigma_{i,2} & \sigma_{i,3} & \sigma_{i,4} & \cdots & 0 \\ \sigma_{i,2} & \sigma_{i,1} & \sigma_{i,0} & \sigma_{i,1} & \sigma_{i,2} & \sigma_{i,3} & \cdots & 0 \\ \sigma_{i,3} & \sigma_{i,2} & \sigma_{i,1} & \sigma_{i,0} & \sigma_{i,1} & \sigma_{i,2} & \cdots & 0 \\ \sigma_{i,4} & \sigma_{i,3} & \sigma_{i,2} & \sigma_{i,1} & \sigma_{i,0} & \sigma_{i,1} & \cdots & 0 \\ 0 & \sigma_{i,4} & \sigma_{i,3} & \sigma_{i,2} & \sigma_{i,1} & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \sigma_{i,1} \\ 0 & 0 & 0 & 0 & 0 & \cdots & \sigma_{i,1} & \sigma_{i,0} \end{bmatrix} \quad (8)$$

$$B = \begin{bmatrix} \delta_0 & \delta_1 & \delta_2 & \delta_3 & \delta_4 & 0 & \cdots & 0 \\ \delta_1 & \delta_0 & \delta_1 & \delta_2 & \delta_3 & \delta_4 & \cdots & 0 \\ \delta_2 & \delta_1 & \delta_0 & \delta_1 & \delta_2 & \delta_3 & \cdots & 0 \\ \delta_3 & \delta_2 & \delta_1 & \delta_0 & \delta_1 & \delta_2 & \cdots & 0 \\ \delta_4 & \delta_3 & \delta_2 & \delta_1 & \delta_0 & \delta_1 & \cdots & 0 \\ 0 & \delta_4 & \delta_3 & \delta_2 & \delta_1 & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \delta_1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & \delta_1 & \delta_0 \end{bmatrix}. \quad (9)$$

Using (7), it is then possible to derive the GMM estimates for α and β as described in section 2.2 and to draw inference on the values of these parameters that is robust even in the presence of aggregate

shocks and individual-specific error variances. The elements of (8) and (9) are then computed as in (10)–(12).

$$\sigma_{i,0} = T^{-1} \sum_{t=1}^T \varepsilon_{i,t}^2 \quad (10)$$

$$\sigma_{i,|l-m|} = [T - |l - m|]^{-1} \sum_{t=1}^{T-|l-m|} \varepsilon_{i,t} \varepsilon_{i,t+|l-m|} \quad (11)$$

for $|l - m| = 0, \dots, \tau$

$$\delta_{|l-m|} = [(N - 1)(T - |l - m|)]^{-1} \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{t=1}^{T-|l-m|} \varepsilon_{i,t} \varepsilon_{j,t+|l-m|} \quad (12)$$

for $|l - m| = 0, \dots, \tau$

Appendix 2. Sensitivity to Choice of Event Thresholds

Our analysis of the individual and aggregate SPF probability forecasts for fixed threshold events has been based on upper and lower thresholds of 1 percent and 2 percent. The choice of these thresholds was primarily motivated by their interest from a user perspective. For example, inflation above 2 percent is seen as inconsistent with the ECB definition of price stability, while rates that are below 1 percent are often seen as raising concerns about disinflationary or even deflationary pressure. Notwithstanding the interest in these thresholds from an economic perspective, it is interesting to assess the sensitivity of our broad conclusions to the precise numerical choice of threshold values. Such a sensitivity analysis is conducted for the aggregate- and individual-level analysis in tables 10 and 11, respectively. In particular, these tables report the p-values corresponding to the tests of well-calibrated forecasts and positive resolution whilst varying the upper threshold over the interval [1.75, 2.00, 2.25] and the lower threshold over the interval [0.75, 1.0, 1.25]. Overall, the sensitivity analysis reported in tables 10 and 11 highlights that our main results appear quite stable when we vary the numerical thresholds for the fixed threshold events. At the 10 percent level, rarely do we observe a change in the result of the hypotheses test when the

Table 10. Sensitivity of Inference to Variation of Event Thresholds:
Aggregate Forecast (event threshold $+/- 0.25\%$)

<i>A. Event 1 (Target Variable > x%)</i>						
	<i>x = 1.75</i>		<i>x = 2.0</i>		<i>x = 2.25</i>	
	$H_0:$	$H_0:$	$H_0:$	$H_0:$	$H_0:$	$H_0:$
	$\alpha = 0, \beta = 1$	$\beta \leq 0$	$\alpha = 0, \beta = 1$	$\beta \leq 0$	$\alpha = 0, \beta = 1$	$\beta \leq 0$
	GDP Growth > x% (H = 1)	0.51	0.00	0.33	0.00	0.42
	GDP Growth > x% (H = 2)	0.00	0.74	0.00	0.79	0.00
	Inflation > x% (H = 1)	0.00	0.13	0.00	0.61	0.01
	Inflation > x% (H = 2)	0.00	0.99	0.00	1.00	0.00
<i>B. Event 2 (Target Variable < x%)</i>						
	<i>x = 0.75</i>		<i>x = 1.0</i>		<i>x = 1.25</i>	
	$H_0:$	$H_0:$	$H_0:$	$H_0:$	$H_0:$	$H_0:$
	$\alpha = 0, \beta = 1$	$\beta \leq 0$	$\alpha = 0, \beta = 1$	$\beta \leq 0$	$\alpha = 0, \beta = 1$	$\beta \leq 0$
	GDP Growth < x% (H = 1)	0.22	0.00	0.24	0.00	0.44
	GDP Growth < x% (H = 2)	0.07	0.62	0.04	0.66	0.01
	Inflation < x% (H = 1)	0.02	0.30	0.00	0.60	0.06
	Inflation < x% (H = 2)	0.00	0.92	0.00	0.93	0.00

Table 11. Sensitivity of Inference to Variation of Event Thresholds:
Pooled Sample (event threshold $+/- 0.25\%$)

<i>A. Event 1 (Target Variable > x%)</i>						
	<i>x = 1.75</i>		<i>x = 2.0</i>		<i>x = 2.25</i>	
	H₀:	H₀:	H₀:	H₀:	H₀:	H₀:
	$\alpha = 0, \beta = 1$	$\beta \leq 0$	$\alpha = 0, \beta = 1$	$\beta \leq 0$	$\alpha = 0, \beta = 1$	$\beta \leq 0$
	GDP Growth > x% (H = 1)	0.10	0.01	0.05	0.01	0.12
	GDP Growth > x% (H = 2)	0.00	0.74	0.00	0.80	0.00
Inflation > x% (H = 1)	0.00	0.02	0.00	0.00	0.00	0.24
	Inflation > x% (H = 2)	0.00	0.99	0.00	1.00	0.00
<i>B. Event 2 (Target Variable < x%)</i>						
	<i>x = 0.75</i>		<i>x = 1.0</i>		<i>x = 1.25</i>	
	H₀:	H₀:	H₀:	H₀:	H₀:	H₀:
	$\alpha = 0, \beta = 1$	$\beta \leq 0$	$\alpha = 0, \beta = 1$	$\beta \leq 0$	$\alpha = 0, \beta = 1$	$\beta \leq 0$
	GDP Growth < x% (H = 1)	0.14	0.00	0.14	0.00	0.20
	GDP Growth < x% (H = 2)	0.00	0.63	0.00	0.69	0.00
Inflation < x% (H = 1)	0.01	0.21	0.00	0.35	0.00	0.17
Inflation < x% (H = 2)	0.00	0.88	0.00	0.91	0.00	0.94

thresholds are varied in this way. One exception is that the hypothesis of a negative β tends to be rejected when the threshold for the high inflation outcome ($H = 1$) is lowered from 2 percent to 1.75 percent in the panel regressions (table 11). Lowering the threshold for this event thus casts the inflation probability forecasts in a somewhat better light. At the same time, however, it can be observed from table 11, panel A, that the hypothesis that these probability forecasts are perfectly calibrated is rejected for all three fixed thresholds.

References

- Andrade, P., E. Ghysels, and J. Idier. 2011. "Tails of Inflation Forecasts and Tales of Monetary Policy." Mimeo, Banque de France.
- Benjamini, Y., and Y. Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B* 57 (1): 289–300.
- Berkowitz, J. 2001. "Testing Density Forecasts with Applications to Risk Management." *Journal of Business and Economic Statistics* 19 (4): 465–74.
- Boero, G., J. Smith, and K. F. Wallis. 2008. "Evaluating a Three-Dimensional Panel of Point Forecasts: The Bank of England Survey of External Forecasters." *International Journal of Forecasting* 24 (3): 354–67.
- . 2011. "Scoring Rules and Survey Density Forecasts." *International Journal of Forecasting* 27 (2): 379–93.
- Bowles, C., R. Friz, V. Genre, G. Kenny, A. Meyler, and T. Rautanen. 2010. "An Evaluation of the Growth and Unemployment Forecasts in the ECB SPF." *Journal of Business Cycle Measurement and Analysis* 2010 (2): 63–90.
- Brier, G. W. 1950. "Verification of Forecasts Expressed in Terms of Probability." *Monthly Weather Review* 78 (1): 1–3.
- Christoffersen, P. F. 1998. "Evaluating Interval Forecasts." *International Economic Review* 39 (4): 841–62.
- Clements, M. P. 2006. "Evaluating the Survey of Professional Forecasters Probability Distributions of Expected Inflation Based on the Derived Event Probability Forecasts." *Empirical Economics* 31 (1): 49–64.

- . 2010. "Explanations of Inconsistencies in Survey Respondents' Forecasts." *European Economic Review* 54 (4): 536–49.
- Clements, M. P., F. Joutz, and H. O. Stekler. 2007. "An Evaluation of the Forecasts of the Federal Reserve: A Pooled Approach." *Journal of Applied Econometrics* 22 (1): 121–36.
- Conflitti, C., C. De Mol, and D. Giannone. 2012. "Optimal Combination of Survey Forecasts." ECARES Working Paper No. 2012-023, Université Libre de Bruxelles.
- Croushore, D. 2010. "An Evaluation of Inflation Forecasts from Surveys Using Real-Time Data." *BE Journal of Macroeconomics: Contributions* 10 (1, Article 10): 1–32.
- D'Agostino, A., K. McQuinn, and K. Whelan. 2012. "Are Some Forecasters Really Better Than Others?" *Journal of Money, Credit and Banking* 44 (4): 715–32.
- Davies, A., and K. Lahiri. 1995. "A New Framework for Analyzing Survey Forecasts Using Three-Dimensional Panel Data." *Journal of Econometrics* 68 (1): 205–27.
- Dawid, A. P. 1984. "Statistical Theory: The Prequential Approach." *Journal of the Royal Statistical Society: Series A* 147 (2): 278–90.
- Diebold, F. X., and G. D. Rudebusch. 1989. "Scoring the Leading Indicators." *Journal of Business* 62 (3): 369–91.
- Diebold, F. X., A. S. Tay, and K. F. Wallis. 1999. "Evaluating Density Forecasts of Inflation: The Survey of Professional Forecasters." In *Cointegration, Causality and Forecasting: A Festschrift in Honour of Clive W. J. Granger*, ed. R. Engle and H. White, 76–90 (chapter 3). Oxford: Oxford University Press.
- Engelberg, J., C. F. Manski, and J. Williams. 2009. "Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters." *Journal of Business and Economic Statistics* 27 (1): 30–41.
- Galbraith, J. W., and S. van Norden. 2011. "Kernel-Based Calibration Diagnostics for Inflation and Recession Probability Forecasts." *International Journal of Forecasting* 27 (4): 1041–57.
- . 2012. "Assessing Gross Domestic Product and Inflation Probability Forecasts Derived from Bank of England Fan Charts." *Journal of the Royal Statistical Society: Series A* 175 (3): 713–27.
- Garcia, J. A. 2003. "An Introduction to the ECB's Survey of Professional Forecasters." ECB Occasional Paper No. 8 (September).

- Genre, V., G. Kenny, A. Meyler, and A. Timmermann. 2013. "Combining the Expert Forecasts: Can Anything Beat the Simple Average?" *International Journal of Forecasting* 29 (1): 108–21.
- Giordani, P., and P. Söderlind. 2006. "Is There Evidence of Pessimism and Doubt in Subjective Distributions? Implications for the Equity Premium Puzzle." *Journal of Economic Dynamics and Control* 30 (6): 1027–43.
- Gneiting, T., F. Balabaoui, and A. E. Raftery. 2007. "Probabilistic Forecasts, Calibration and Sharpness." *Journal of the Royal Statistical Society: Series B* 69 (2): 243–68.
- Granger, C. W. J., and M. H. Pesaran. 2000. "Economic and Statistical Measures of Forecast Accuracy." *Journal of Forecasting* 19 (7): 537–60.
- Harding, D., and A. Pagan. 2010. "Can We Predict Recessions?" NCER Working Paper No. 69 (December).
- Holden, K., and D. A. Peel. 1990. "On Testing for Unbiasedness and Efficiency of Forecasts." *The Manchester School* 58 (2): 120–27.
- Hsiao, C. 2003. *Analysis of Panel Data*. Econometric Society Monographs, Book 34. Cambridge University Press.
- Jore, A. S., J. Mitchell, and S. P. Vahey. 2010. "Combining Forecast Densities from VARs with Uncertain Instabilities." *Journal of Applied Econometrics* 25 (4): 621–34.
- Keane, M. P., and D. E. Runkle. 1990. "Testing the Rationality of Price Forecasts: New Evidence from Panel Data." *American Economic Review* 80 (4): 714–35.
- Kenny, G., T. Kostka, and F. Masera. 2014a. "Density Characteristics and Density Forecast Performance: A Panel Analysis." *Empirical Economics* 48 (3): 1203–31.
- . 2014b. "How Informative Are the Subjective Density Forecasts of Macroeconomists?" *Journal of Forecasting* 33 (3): 163–85.
- Lahiri, K., and J. G. Wang. 2006. "Subjective Probability Forecasts for Recessions: Evaluation and Guidelines for Use." *Business Economics* 41 (2): 26–37.
- . 2007. "The Value of Probability Forecasts as Predictors of Economic Downturns." *Applied Economics Letters* 14 (1): 11–14.
- . 2013. "Evaluating Probability Forecasts for GDP Declines Using Alternative Methodologies." *International Journal of Forecasting* 29 (1) 175–90.

- Mincer, J., and V. Zarnowitz. 1969. "The Evaluation of Economic Forecasts." In *Economic Forecasts and Expectations*, ed. J. Mincer. New York: National Bureau of Economic Research.
- Mitchell, J., and K. Wallis. 2011. "Evaluating Density Forecasts: Forecast Combinations, Model Mixtures, Calibration and Sharpness." *Journal of Applied Econometrics* 26 (6): 1023–40.
- Murphy, A. H. 1973. "A New Vector Partition of the Probability Score." *Journal of Applied Meteorology* 12 (4): 595–600.
- . 1988. "Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient." *Monthly Weather Review* 116 (12): 2417–24.
- Murphy, A. H., and E. S. Epstein. 1967. "A Note on Probability Forecasts and 'Hedging'." *Journal of Applied Meteorology* 6 (6): 1002–04.
- Murphy, A. H., and R. L. Winkler. 1992. "Diagnostic Verification of Probability Forecasts." *International Journal of Forecasting* 7 (4): 435–55.
- Newey, W. K., and K. D. West. 1987. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (3): 703–8.
- Pesaran, H. M., and A. Timmermann. 2009. "Testing Dependence among Serially Correlated Multicategory Variables." *Journal of the American Statistical Association* 104 (485): 325–37.
- Tay, A. S., and K. F. Wallis. 2000. "Density Forecasting: A Survey." *Journal of Forecasting* 19 (4): 235–54. Reprinted in *A Companion to Economic Forecasting*, ed. M. P. Clements and D. F. Hendry, 45–68. Oxford: Blackwell, 2002.
- Yates, J. F. 1982. "External Correspondence: Decompositions of the Mean Probability Score." *Organizational Behavior and Human Performance* 30 (1): 132–56.

Monetary Policy Committees: Comparing Theory and “Inside” Information from MPC Members*

Mikael Apel,^a Carl Andreas Claussen,^a
Petra Lennartsdotter,^a and Øistein Røisland^b

^aSveriges Riksbank

^bNorges Bank

How do monetary policy committee (MPC) members form their views about the appropriate level of the interest rate? To what extent do they change their minds during the deliberations in the policy meeting? How important is the chairman? These are issues discussed and analyzed in the literature but so far with limited evidence of how this actually works in practice. We asked former and serving MPC members in Sweden and Norway. The responses confirm a number of theoretical assumptions and predictions, but some results also challenge the conventional view of monetary policymaking by committees. For instance, there seems to be rather limited pooling within the committee; input from the staff is more important than input from other MPC members. Members tend to have decided beforehand how to vote at the policy meeting, and they have a firm idea of what the positions of their colleagues will be. We also find that there is a preference for the status quo regarding institutional features in both central banks, implicitly suggesting that the exact institutional setup is not

*The views expressed herein are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Executive Boards of Norges Bank or Sveriges Riksbank. We thank the members of the Executive Boards of Norges Bank and the Riksbank for responding to the survey, thereby making this study possible. We are also grateful for comments and suggestions by Petra Gerlach-Kristen, the editor, and the anonymous referees. Author e-mails: Mikael.Apel@riksbank.se, Carl-Andreas.Claussen@riksbank.se, Petra.Lennartsdotter@riksbank.se, and Oistein.Roisland@Norges-Bank.no.

crucial. One way of interpreting this would be to say that the two central banks have managed to establish processes that, according to the respondents, ensure the quality of policy decisions irrespective of the institutional setup.

JEL Codes: D71, E52, E58.

1. Introduction

Monetary policy is in many countries set by a committee. The shift away from a single policymaker in charge of interest rate setting has triggered a large number of papers examining the merits and shortcomings of monetary policy committees (MPCs). Most of these papers are theoretical and use models that make assumptions about the characteristics of MPC members, about the way information is gathered and exchanged, and about the procedures and dynamics of the actual policy meetings. Such assumptions influence the results these models yield and the policy recommendations they imply.

In this paper we reexamine assumptions, results, and ambiguous empirical findings in the literature. We do so by asking the people who should know best—actual MPC members. The MPC members interviewed are from Norges Bank (the central bank of Norway) and Sveriges Riksbank (the central bank of Sweden). While both banks operate the same type of flexible inflation targeting in very similar cultures and economies, the Norwegian MPC can be viewed as a collegial committee, whereas the Swedish MPC is generally seen as an individualistic committee.

The research method of posing direct questions to economic agents has traditionally been subject to some skepticism but has become more common during the last two decades.¹ Surveys may add important information on what lies behind agents' decisions, thus enabling better interpretations of findings based on data on realized decisions and potentially pointing to new directions for future theoretical research. For example, studies asking firms about their price-setting behavior have been conducted in a large number of countries, generating valuable information on, for instance,

¹ See Boulier and Goldfarb (1998) for an overview of the methodological debate.

the frequency of price reviews and the sources of price rigidity.² Surveys have also been used to investigate central banks' views as regards monetary policy—for example, why credibility is considered important and what makes a central bank credible (Blinder 2000).

While surveys have limitations, we believe they can provide useful evidence on monetary policymaking by committees, in much the same way as they have in other areas.³ The information contained in our survey is unique, as it is to our knowledge the only survey in the literature where the respondents are actual MPC members. In addition, surveying two different types of MPC in two relatively similar countries with the same type of monetary policy regime provides an interesting comparative analysis.

The issues covered in the paper are the optimal size and composition of an MPC (section 3), the deliberation within MPCs (section 4), the role of the governor (section 5), the heterogeneity between committee members (section 6), and the merits of transparency (section 7). In each section we relate our results to the relevant literature. We discuss some of our results and conclude in the final section of the paper.

Several assumptions and results from the previous literature are confirmed by our survey: a committee is preferable to a single policymaker, but the committee should not be very large; the chairman seems to have more influence in a collegial committee than in an individualistic committee; and MPC members do indeed have heterogeneous preferences.

Other assumptions and results are not confirmed by our survey. There seems, for instance, to be a rather limited amount of pooling of information within the committee. Instead, committee members rely heavily on input from the staff. As regards the policy meeting, members tend to have decided beforehand how to vote and very rarely change their minds during the meeting. They also have a firm idea already before the meeting of their colleagues' positions, and

²Studies include Blinder et al. (1998) for the United States; Fabiani et al. (2006) for the euro area; Hall, Walsh, and Yates (2000) for the United Kingdom; Amirault, Kwan, and Wilkinson (2004) for Canada; Apel, Friberg, and Hallsten (2005) for Sweden; and Langbraaten, Nordbø, and Wulfsberg (2008) for Norway.

³Limitations arise because responses may be sensitive to the precise wording of the questions or because respondents sometimes have incentives to conceal the truth, etc. See Blinder et al. (1998) for a discussion.

this is not only the case for the full-time committee members at the Riksbank but also for the part-time members at Norges Bank. Thus, the monetary policy decision is more likely to be affected by “pooling by voting” than “pooling by talking.”

When asked about the role of external MPC members, the choice of communication strategy, and whether voting records should be published, our respondents displayed a tendency to support the status quo in their respective committee. This indicates that members are content with the present institutional setup and suggests that there is not one obvious preferable choice in this respect.

2. The Questionnaire and the Respondents

2.1 Facts about Sveriges Riksbank and Norges Bank

Both the Riksbank and Norges Bank are inflation targeters, but they differ in certain institutional respects, such as the composition of their monetary policy committees.

The Executive Board of the Riksbank consists of six members. One of the members is appointed governor and is chairman of the Executive Board. Members are appointed for a period of five or six years according to a rolling schedule. There is no restriction on the number of periods for which a member may be reappointed. Several members have served on the board for more than one term, and one member has been reappointed twice. When appointing members, a guiding principle is that “the appropriate competence” should be represented on the Executive Board (Government Bill 1997/98:40). With few exceptions, members have been recruited from outside the bank, typically from the financial sector, government agencies, or academia.

All six Riksbank MPC members are internal full-time members. There is no staff attributed to a particular member, and the members do not hire their own staff. Interest rate decisions are made by voting, with results announced the following day, and the members can publicly express their views. If an equal number of members support different proposed decisions, the chairman has the casting vote. The Riksbank board is thus considered an individualistic, as opposed to a collegial, committee.

The Riksbank is today ranked among the world's most transparent central banks.⁴ The minutes of the monetary policy meeting have been published with a lag of two weeks practically ever since the board was established in 1999. Statements in the minutes have been attributed since May 2007, showing who said what during the discussion. Information about who has dissented and their preferred policy rate was initially published together with the minutes. Since May 2009 this information has also been provided in the press release announcing the interest rate decisions.

In the period from 1999 to 2009 (the years covered in the Riksbank survey), there were dissenting votes at about one-third of the monetary policy meetings. Except for the governors and a member who had only recently been appointed at the time of the survey, all members in the sample had dissented at least once. Disagreement was considerably more common in 2010–13, increasing the average share of meetings with dissents for the whole period to more than a half.

The Executive Board of Norges Bank consists of seven members: the governor, the deputy governor, and five external members. The governor and the deputy governor are appointed for six years and may be reappointed once. The external members are appointed for four years according to a rolling schedule and may be reappointed twice. Several members have been on the board for two terms. When appointing members, importance is given to ensuring a breadth of background and expertise, with particular emphasis on economics and finance and a good grasp of socioeconomic issues (Proposition to the Odelsting No. 81 (2002–2003)). Typically, the external members are recruited from outside the bank and have an academic, political, and/or business background.

The external members of the Executive Board of Norges Bank work part-time and keep their ordinary jobs while serving on the board. There is no staff attributed to a particular member, and the members do not hire their own staff. The board is a collegial committee, where the governor seeks a consensus decision. Formally, each member has one vote. The members may dissent from the governor's interest rate proposal, but dissents are not made public for a long

⁴See, e.g., Dincer and Eichengreen (2013) and Eijffinger and Geraats (2006).

period (sixty years up to March 2012, twelve years since). Attributed minutes from the policy meetings are not provided, but since 2011 a summary of the Executive Board's assessment has been published in each Monetary Policy Report.

Table 1 shows that Norges Bank and the Riksbank have smaller MPCs and less frequent monetary policy meetings than the Bank of England (BoE), the Federal Reserve (Federal Open Market Committee, or FOMC), and the European Central Bank (ECB). At these three central banks, MPC members have their own dedicated staff, unlike at Norges Bank and the Riksbank. Dissents on the Swedish MPC are more frequent than on the Federal Open Market Committee but slightly less common than on the Bank of England's MPC.

Figure 1 shows the five-year moving average of headline inflation and the inflation targets in the two countries. Except for some years around 2000, the paths of inflation have been quite similar. Policy rate movements have also been quite similar; see figure 2. During the period 1999–2013 Norges Bank changed the policy rate fifty times, and the Riksbank forty-eight times.

2.2 *The Survey*

Most questions in the survey focused on institutional design, the decision-making process, and communication. Typically, members were asked to grade a statement on a four-grade scale, e.g., “unimportant,” “not very important,” “fairly important,” and “very important.” In order to facilitate comparison, we phrased the questions to the members of the two MPCs as similarly as possible. However, due to institutional differences, some questions were somewhat adjusted. Respondents had the opportunity to provide comments in connection with each block of questions. For some issues we also explicitly asked the members if they had other explanations or views than the ones suggested in the questionnaire.

The main survey was conducted during the period May–June 2009 for the Riksbank MPC members and during the period March–April 2011 for the Norwegian respondents. We sent the questionnaire to present and former members of the Riksbank's and Norges

Table 1. MPC Features and Inflation (2000–13)

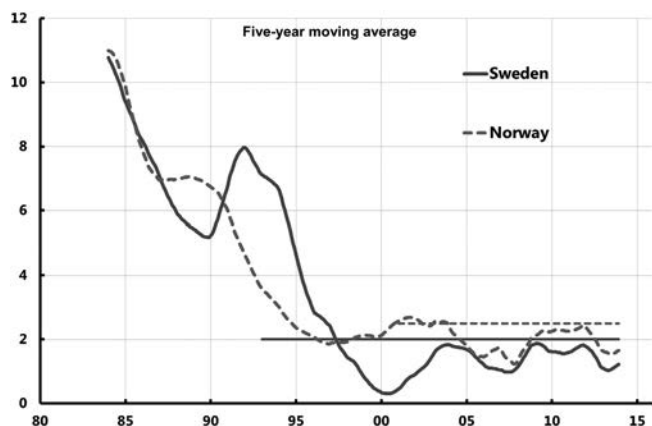
	MPC				Inflation ^a		
	Size	Meetings per Year	Meetings with Dissents ^a	Individualistic	Own Staff	Objective	Mean
Norges Bank	7	6	N/A	No	No	2.5	2.0
Sveriges Riksbank	6	6	51	Yes	No	2	1.4
FOMC	12	8	34*	No	Yes	N/A	2.4
BoE	9	12	61*	Yes	Yes	2/(2.5)**	2.2
ECB	23	12	N/A	No	Yes	≤ 2	2.1

^aPercent.

*Includes votes on asset purchases.

**The inflation objective for the Bank of England was changed in December 2003 from 2.5 percent RPIX to 2 percent CPI inflation.

Figure 1. CPI Inflation and Inflation Targets in Norway and Sweden

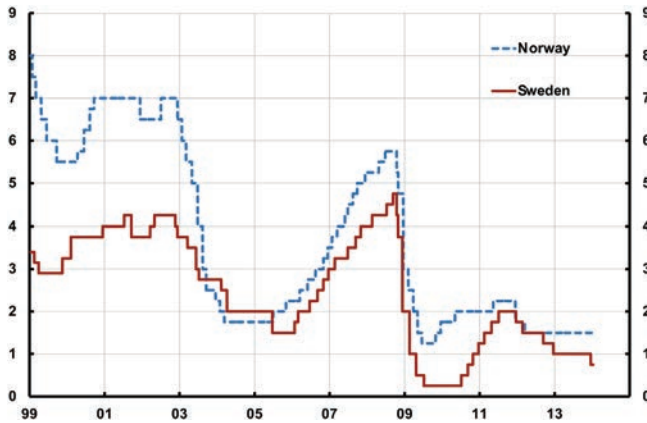


Bank's MPCs in the period from 1999 until the date of the survey. A follow-up survey was sent to some of the members in 2013; see below. Because of the small number of internal members (governors and deputy governors) at Norges Bank in the period we considered, it was difficult to ensure their anonymity. We therefore decided to ask only external members. Since 1999 and up to the date of the survey, a total of thirteen people served on the Riksbank's MPC, while seventeen external members served on Norges Bank's MPC.⁵ In order to maximize the number of responses and encourage honesty, we emphasized the confidentiality of individual responses. From the Riksbank we received answers from twelve out of thirteen MPC members, while from Norges Bank we received answers from thirteen out of seventeen members.

In October 2013 we conducted a smaller follow-up survey, with the purpose of obtaining clarifying information about certain issues in the original study. The survey contained supplementary questions on the influence of fellow colleagues on the members' own decisions,

⁵We have included two alternate members of Norges Bank's MPC, because from 2010 the alternate members have attended the meetings and the discussions in the same way as the ordinary members, except that they have no formal voting powers.

Figure 2. Policy Rate Developments in Norway and Sweden



the perceived pros and cons of committee decision making during a crisis, and views on internal vs. external members. As we did not deem it possible to approach all the previous respondents again, we sent the follow-up survey to the six members of the Riksbank's MPC of 2009 and to four members of Norges Bank's MPC of 2011. We received responses from all ten. In the following, the results from this survey are discussed where relevant.

We next turn to the first set of questions in the survey, which deal with the size and composition of the MPC. For expositional reasons, we have chosen to report only the median responses for most of the questions, but report or comment on the dispersion when this information is of particular interest or when the distribution of responses is much skewed.

3. Size and Composition

3.1 *Why an MPC?*

Blinder (2006) argued that MPCs were not necessary before central banks became independent because “there is no reason to have a monetary policy committee when the central bank is simply taking

orders from its government” (p. 21). With independence came the trend towards MPCs.

In the literature, the main rationale for having a group of people, rather than one individual, responsible for monetary policy decisions is that “several heads are better than one.” The idea is that MPCs improve on decisions by pooling members’ information and knowledge (see, e.g., Blinder 2007).⁶ Moreover, pooling is an insurance against extreme preferences any single MPC member may hold.

The perception that groups tend to make better monetary policy decisions than individuals has received empirical support. For example, Blinder and Morgan (2005, 2008a, 2008b) and Lombardelli, Proudman, and Talbot (2005) find in experiments where students make monetary policy decisions that the decisions are better if made by a group. In a theoretical study, Gerlach-Kristen (2006) finds that if perceptions of the state of the economy differ between individuals, a committee achieves better policy outcomes than a single policymaker.

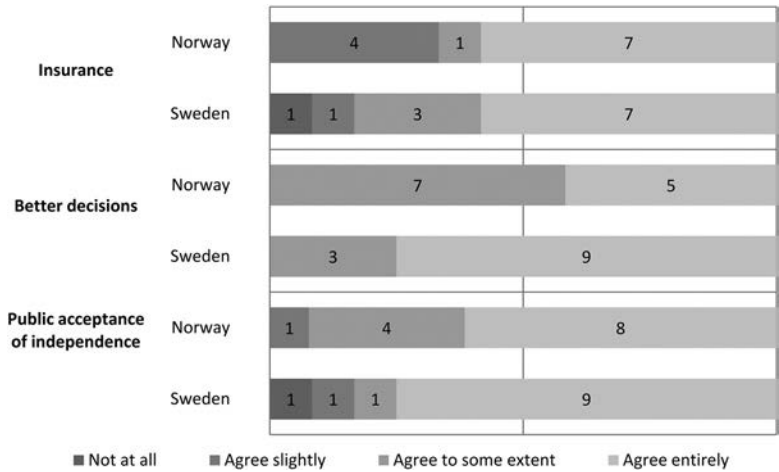
What are the views of actual MPC members on the reasons for having a committee? We asked the members to what extent they agreed with the following statements:

- *Public Acceptance of Independence:* Sufficient acceptance among the general public and politicians of a system where the conduct of monetary policy is delegated to an (independent) central bank requires that monetary policy decisions are made by a group of people and not by one individual.
- *Better Decisions:* Monetary policy decisions will be better if they are made by a group of people instead of one individual.
- *Insurance:* Making decisions as a group functions as an insurance against extreme preferences one individual governing monetary policy might hold.

Figure 3 shows the distribution of responses, where the vertical line represents the median response. We see that the majority of the

⁶This finding goes back to Condorcet’s jury theorem, which dealt with court decisions.

Figure 3. Why Monetary Policy by Committee?



Swedish and Norwegian respondents agreed with all three reasons for having a committee.⁷

Are there circumstances where committee decision making is perceived as less favorable? Such a situation might arise when decisions have to be made very quickly, like during a crisis. This is an issue we addressed in the follow-up survey. We asked the following questions: If you compare the situation during a crisis with a more normal situation, what pros and cons do you see with decisions in the crisis situation being made by a group of people? Do the pros outweigh the cons or vice versa?

All respondents, in both Norway and Sweden, found that the pros of group decision making outweigh the cons even in a crisis. Members found that groups are more likely to come up with sufficient information and all relevant arguments, that groups reduce the risk of making large mistakes, and that a group can better lend legitimacy and credibility to serious decisions. Some members mentioned the potential shortcoming that a group may not be able to make decisions sufficiently fast. It was, however, not given much weight, as it

⁷In some cases, like here, some respondents did not answer all questions. In this case the median is not affected by the two missing answers.

was considered not to be borne out by practical experience. Several members of the Swedish MPC stressed that there were no problems of this kind during the financial crisis. One respondent raised the possibility that considerable disagreement within the group may damage public confidence, which is particularly problematic during a crisis. Others argued that decisions tend to be easier to agree upon during a crisis, implicitly suggesting that such a situation is not very likely.

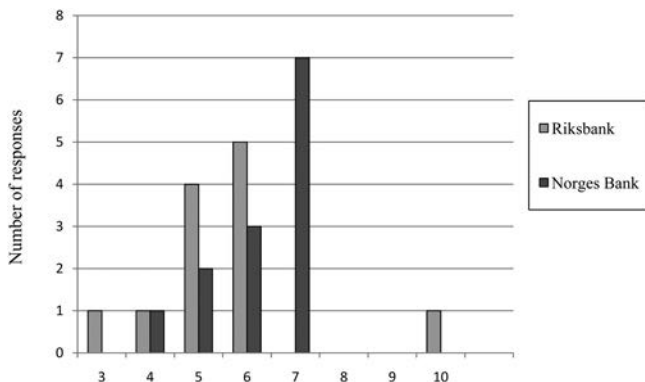
3.2 The Optimal Size

If MPCs improve decisions by pooling members' information and knowledge, this suggests in isolation that the bigger the committee, the better. But there are also costs associated with increasing the size of a committee, such as less thorough (or more time-consuming) discussions and more potential for free-riding, i.e., preparing poorly for policy meetings and instead relying on input from other MPC members.

The question of the optimal MPC size is addressed directly or indirectly in different strands of the literature. In an experimental study on monetary policy decisions, Blinder and Morgan (2008b) find that eight-person groups do not perform significantly better than four-person groups. In a review of the rather small economics literature on committees as well as the older and larger literature on committees in other social sciences, Sibert (2006) concludes that the ideal MPC should not have more than five members. In another interdisciplinary survey of studies, Erhart and Vasquez-Paz (2007) conclude that the optimal committee consists of roughly five to nine members.⁸ In simulations using a theoretical model of dynamic decision making by a monetary policy committee, Weber (2010) arrives at the conclusion that there is relatively little benefit in having a committee with more than seven members.

We therefore asked the members what they believed to be the most appropriate size of their respective MPCs. When asked if they

⁸Gerling et al. (2005) survey the game theory literature on committee decision making but do not settle on a specific figure. Berger and Nitsch (2008) provide an international comparison of MPC size.

Figure 4. Appropriate Size of MPC

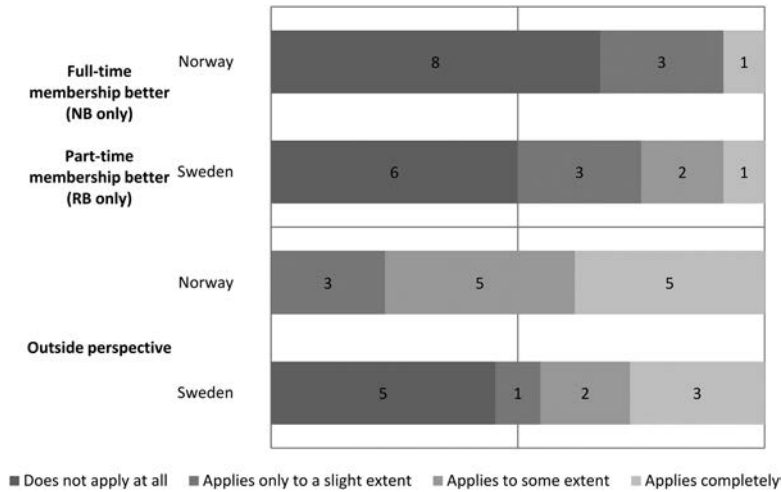
thought their current six-member board was too small, the right size, or too large, half of the Riksbank MPC members thought it was too large, and only one member thought it was too small. The Norwegian respondents were more satisfied with their own seven-member board. More than two-thirds of the members thought that the size was right, while the rest thought it was too large. When asked to state their preferred committee size, the responses were distributed as shown in figure 4. Although the preferred size is within the estimated span of the optimal size in the literature, the average response lies within the lower region of that span.

In the space provided for comments, one respondent wrote that finding qualified people without conflicts of interests can be a problem in a small country like Norway or Sweden. No other member gave any explanation as to why they preferred a specific size. However, as we shall see below, members do not consider input from other MPC members very important for their own views, and this may explain why they believe a relatively small committee is sufficient.

3.3 The Value of Outsiders

One aspect of the composition of an MPC that has received much attention in the literature is the inclusion of external members. One could argue that MPC members should be heterogeneous in order to prevent a high positive correlation in individual judgment errors.

Figure 5. Internal/Full-Time vs. External/Part-Time Members



Related to this, Sibert (2006) argues that external members help avoid “groupthink,” which occurs when the committee members’ striving for consensus makes them stop paying sufficient attention to alternative courses of action. The committee may then make worse decisions than any of its members would have made on their own.

To investigate views on the issue of internal vs. external members, we asked the Riksbank MPC members to what extent they thought the following statement applied: Having external members would contribute to better decisions, as it adds information, knowledge, and an “outside perspective.” The Norges Bank MPC members were asked the same question, but “would contribute” was replaced with “contributes.”

The responses are shown in figure 5. The Riksbank MPC members tend not to support the view that having external members would provide a useful “outside perspective” on decisions, although the dispersion in the responses was quite large. One of the members wrote that an MPC with personal responsibility, as in Sweden, should consist of members who all have the same formal status: “Otherwise, they do not have the same possibility to prepare for the decisions to be made.” The same member thought the “outside

perspective” can be ensured by members reading external analyses and talking with other people active in various areas. Another member found that “it is more important to have different skills represented in the staff and that the MPC is composed of members with skills acquired from work outside the bank.” The Norwegian respondents, however, quite strongly support the view that having external members is beneficial.

The external vs. internal issue was one of the topics addressed in the follow-up survey conducted in October 2013. Even though we had already obtained some explanations from the respondents in the original survey, we wanted to know more about the motives behind the MPCs in both Norway and Sweden defending their respective setup. In the follow-up survey, we asked the following: Assume that you were to set up a monetary policy committee “from scratch.” Would you recommend that the committee consists of only internal members, or should it also include external members who are not employed full-time at the central bank but have their main occupation elsewhere, for example in public administration or academia? What are the main motives for your recommendation?

The responses essentially confirmed the results from the original survey. Only one of the six Swedish respondents advocated external members, thinking it would reduce the risk of groupthink. Another respondent was somewhat inconclusive, basically listing pros and cons. This member argued that if the committee is in charge of the bank’s complete set of activities, as is the case today, “it would perhaps be more natural if it were internal.” The Swedish respondents favoring only internal members believed it would be difficult to have a judicious and focused forecasting and decision-making process with a committee containing external members, and that internal members are likely to have an informational advantage that makes it harder for external members to influence the decisions. Members also found that in a small country like Sweden it may be difficult to find external members for whom there is no conflict of interest between monetary policy and their regular employment. The Norwegian respondents all favored external members and again gave considerable weight to the groupthink argument.

There was thus a tendency for MPC members to defend their own respective systems and thus the “status quo.” We also found

a preference for the “status quo” in other matters, which we will return to later. In the discussion at the end of the paper we propose some “institutional” explanations for these preferences for the status quo.

To sum up this section, MPC members believe that they make better decisions as a group than as individuals. This is in line with mainstream theory. They also prefer relatively small committees. Their views on external members differ, however. Riksbank MPC members prefer an MPC with only internal members, and the Norges Bank MPC members prefer a committee which also includes external members. These preferences may reflect differences in the practices and prevailing institutional frameworks in the two countries. We will discuss potential institutional explanations for these diverging results in more detail in the last section of the paper.

Having reviewed the composition of MPCs, we next turn to the question of how interest rate decisions are arrived at.

4. Deliberation

An MPC meeting typically consists of two principal stages. In the first stage—the deliberation stage—the members exchange views. In the second stage—the decision stage—the MPC reaches the policy decision through an agreed aggregation rule (majority voting, for example).

Like any economic agents, MPC members receive information, interpret the information, and update their judgments. The theoretical literature often thinks of this process as Bayesian updating: the MPC members have some priors, or “beliefs,” that are confronted with new information, and they form their posterior judgments (Chappell, McGregor, and Vermilyea 2012). In this section, we consider two types of deliberations: the deliberations within the MPC and the deliberations between the MPC and the staff.

At both central banks, the policy process starts at preliminary meetings a few weeks before the final policy meeting; see, e.g., Hallsten and Tägtström (2009) for the case of Sweden. It is therefore legitimate to suspect that members have more or less formed their final and definitive view before the final monetary policy meeting. We end this section by asking whether this is actually the case.

4.1 *Deliberations within the MPC*

In the existing theoretical literature on MPCs, it is assumed that there is scope for committee members to pool their different information sets, which improves policy outcomes; see, e.g., Berk and Bierut (2011), Gerlach-Kristen (2006 and 2008), and Weber (2010). To investigate the relevance of this assumption for MPCs in practice, we asked the respondents how important they considered the other MPC members as a source of information on the following:⁹

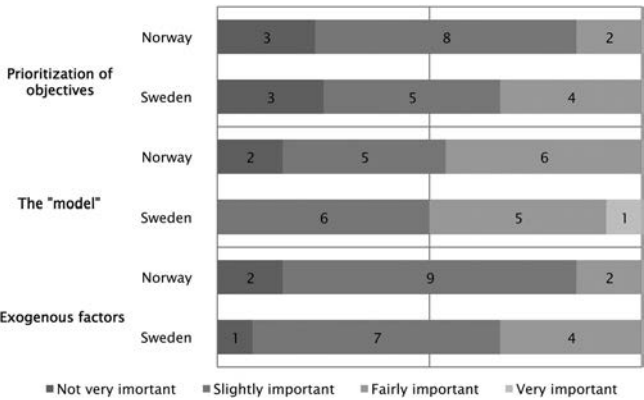
- *Exogenous Factors*: The current economic situation and trends that the bank is unable to influence, such as future oil prices and international economic developments.
- *The “Model”*: How the economy functions and thereby how things will develop if the central bank acts in one particular way or another.
- *Prioritization of Objectives*: How quickly inflation should be brought back on target/prioritizing between stabilizing inflation and stabilizing the real economy.

The results are summarized in figure 6. Most members did not consider their fellow members as very important information sources. The median respondent tended to view other members as “not very important” sources of information on all three issues. One explanation for this is probably that in neither of the two banks do members of the MPCs have their own staff. Most of the work on collecting new information, analyzing this information, doing research, building models, etc. is conducted by the bank’s staff. Indeed, the important role of the staff is confirmed in the next subsection.

In the follow-up study conducted in October 2013, we looked further into this issue. Instead of referring to the three information categories used in the first survey, we posed the question in a general form: How useful would you say your colleagues on the board are/were to you in the process of reaching your monetary policy position? The respondents were asked to rate their colleagues’ contribution as “not very useful,” “somewhat useful,” “quite useful,”

⁹The Norges Bank MPC members were asked how important the other external members were, as the two internal members (governor and deputy governor) also represent the staff.

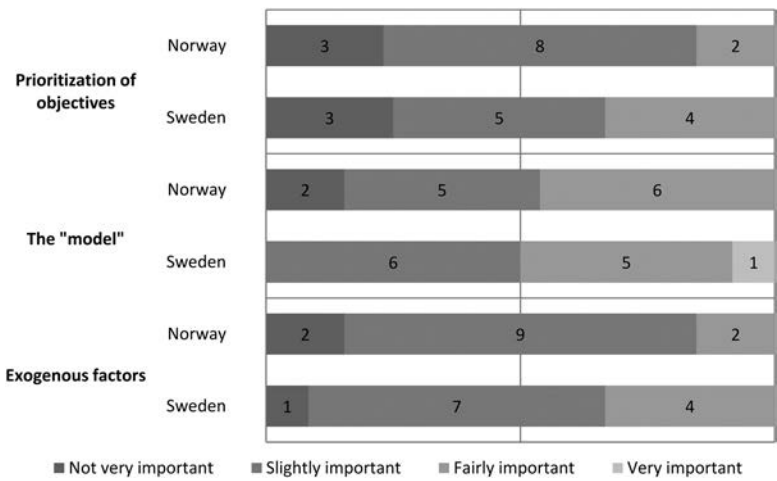
Figure 6. The Importance of Peer Members as Sources of Information



or “very useful.” We also asked the respondents to qualify their answers and, in particular, whether there is any kind of input that the members get from their colleagues that they do not get from the staff.

Most respondents in both Norway and Sweden recognized that there are indeed inputs provided only by their colleagues, such as anecdotal information from a particular field of expertise, different interpretations of the information from the staff, and reports from international meetings. Yet at both Norges Bank and the Riksbank, colleagues seem to play quite a modest role when members form their views. The median response for both banks was that colleagues were only “somewhat useful.” At Norges Bank, this was the alternative chosen by all the respondents, whereas there was some dispersion in the Riksbank responses, with one responding “quite useful” and one “very useful.” One of the respondents from Norges Bank emphasized that the main purpose of having external members is not to provide more information but to ensure that the staff’s arguments are well founded. The same respondent also mentioned that it is advantageous that there are more people to decide on the balance between different objectives, which is partly “political” in nature. Our overall interpretation of the results is that members see the benefits of committee decision making as being mainly about providing a

Figure 7. The Importance of Peer Members as Sources of Information



check-and-balance mechanism, and that “pooling by voting” is more important than “pooling by talking.”

4.2 The Staff’s Role

The theoretical literature on MPCs barely considers the staff’s role in the pooling process.¹⁰ But monetary policy decisions are not made in a vacuum. As noted above, the MPC receives considerable inputs from the central bank staff.

How important are these inputs to MPC members? We asked them the same questions as in section 4.1, this time focusing on the role of the staff. Figure 7 shows the results. Most respondents found the staff more important than their fellow members. This suggests that the existing theoretical literature on MPCs, which focuses on information sharing among the MPC members themselves, could miss an important aspect of the decision-making process. Future theoretical work on collective decision making in central banks would

¹⁰To our knowledge, the only exception is Claussen et al. (2012), who assume that the chairman has access to the staff and is thereby better informed than the other members. They also consider information pooling within the staff prior to the staff advice to the chairman.

therefore benefit from focusing more on the interaction between the staff and the MPC members.¹¹

The staff seem to have more influence regarding exogenous factors and the model (which together form the forecasts) than regarding the prioritization of objectives. This is likely to reflect that the latter represents true *policy* decisions as opposed to *economic analysis*. It might also reflect the staff participating less actively in discussions on how to balance the various policy trade-offs.

Although the results from this survey may not necessarily carry over to other central banks, they could be interpreted as giving some support to the results in Romer and Romer (2008), who found that the Federal Open Market Committee (FOMC) members' forecasts did not add any extra value to the staff forecasts. Based on this, Christina and David Romer argued that "a more effective division of labor within the Federal Reserve System might be for the staff to present policymakers with policy options and related forecast outcomes, and for policymakers to take those forecasts as given. With this division, the role of the FOMC would be to choose among the suggested alternatives, not to debate the likely outcome of a given policy." Our results indicate that to some extent such a division of labor takes place, albeit perhaps implicitly, during the monetary policy process preceding the MPC meetings in Norway and Sweden.

Even though respondents from both central banks considered the staff more important than their fellow MPC members, there were some small differences between the groups. The staff seems even more important for the Riksbank MPC members than for the Norwegian respondents, at least as sources of information on current developments and the workings of the economy. This is likely to reflect that the Riksbank MPC members are internal and have easier access to the staff than the external Norges Bank MPC members.

4.3 *How Much Is Settled before the Meeting?*

As noted above, the policy process in both banks has been under way for a number of weeks by the time of the policy meeting. To

¹¹This issue seems particularly important for central banks that publish interest rate paths, as one of the Riksbank respondents wrote that the role of the staff has increased since publication began. See also the discussion in the conclusion.

find out more about the state of affairs before the meeting, we asked the members the following questions:

- *Decided Prior to Meeting:* How often has it happened that you have in principle decided [how you intend to vote]/[which decision you prefer]¹² before the monetary policy meeting and not changed your mind during the meeting?
- *New Information:* How often has it happened that information and arguments that are new to you have been presented at the monetary policy meeting?
- *Changed My Mind:* How often has it happened that you have changed your mind during the monetary policy meeting despite having in principle decided in advance?
- *Not Decided Prior to Meeting:* How often has it happened that you have not known [how you would vote]/[which decision you would prefer] prior to the monetary policy meeting?

We also wanted to obtain information about how much the members knew about the views of the other members before the meeting. Due to the institutional differences, we asked the two groups of MPC members somewhat different but related questions. The Riksbank MPC members were asked the following:

- *Clear Impression of Majority:* How often has it happened that you have had a clear impression of how the *majority* of your colleagues intend to vote before going into the monetary policy meeting?
- *Clear Impression of All Members:* How often has it happened that you have had a clear impression of how *all* of your colleagues intend to vote before going into the monetary policy meeting?

The Norges Bank MPC members were asked the following:

- *Clear Impression of Other External Members:* How often has it happened that you have had a clear impression of which

¹²The Riksbank MPC members were given the first formulation, and the Norges Bank MPC members were given the second, as there is less explicit voting in the Norwegian collegial MPC.

decision the other external members would prefer before going into the monetary policy meeting?

- *Clear Impression of the Internal Members:* How often has it happened that you have had a clear impression of what interest rate decision the governor (or deputy governor) would propose at the monetary policy meeting?

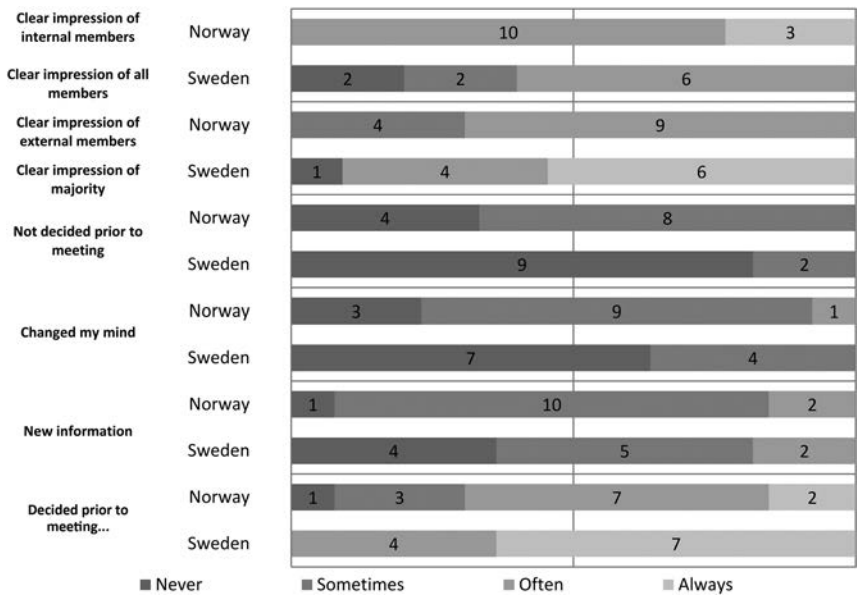
The members' responses are shown in figure 8.¹³ The results show that members have often or always decided beforehand which outcome they favor. This tendency is somewhat stronger for the Riksbank than for the Norges Bank MPC members, as the median Swedish respondent has "always" decided in advance how to vote, while the median Norwegian respondent has "often" decided in advance which level of the policy interest rate he or she would prefer. It is also interesting to note that four Norges Bank MPC members responded that they have "never" or only "sometimes" decided in advance which decision they preferred and not changed their minds, while no one from the Riksbank gave such a response.

This is not surprising. The individual accountability and the detailed attributed minutes from the Riksbank meeting—which are published and which we discuss in detail in section 7 below—give a strong incentive to go to the meeting with a clear view on the appropriate decision. At Norges Bank there are no such incentives. Moreover, Swedish MPC members, being full-time staff, simply may have had more time to make up their minds.

In both committees, it is quite unusual that new information or arguments come up at the monetary policy meeting, and it is even more unusual that members change their minds as a consequence. The median Swedish respondent has "never" changed his or her mind, while the median Norwegian respondent has "sometimes" changed his or her mind. One Norges Bank MPC member has, however, "often" changed his or her mind. This confirms that the Riksbank MPC members have a clearer opinion before the meeting. As stated by Svensson (2009, p. 26): "The discussion and exchange at the final monetary policy meeting do not start from scratch, but are the culmination and summary of [a long series of] meetings. Therefore, one would not expect too much spontaneity

¹³One Swedish respondent did not answer these questions.

Figure 8. The Situation before the Meeting



but rather the presentation of the essential summaries and the reasons for the decision by each member.” At the Riksbank there is even a meeting a few weeks before the final monetary policy meeting where “the Executive Board attempts to arrive at a repo rate path that it seems likely that the majority of the board members can support” (Hallsten and Tägtström 2009, p. 82).¹⁴ At Norges Bank the process leading to the publication of the Monetary Policy Report also includes several meetings and seminars with board members.

Another finding is that the majority of the members of both MPCs usually have a pretty good impression about the decisions

¹⁴The main reason for this is that a Monetary Policy Report (or a shorter Monetary Policy Update) explaining the policy decision is published at the same time as the decision is announced. It should, however, be emphasized that although the members of the boards may have formed an opinion during the process, the final decision on the repo rate is made at the policy meeting. The process is designed so that the forecasts and the Monetary Policy Report can be changed following the meeting.

preferred by the other members. The median Riksbank MPC member responded that he or she “always” had a clear impression of what the outcome of the vote would be. The Norges Bank MPC members, who meet less frequently due to their part-time membership, also tend to have a clear picture of which decision the other external members, as well as the governor, would prefer. This result is not so surprising given the process of “preparatory” meetings described above.

To sum up this section, it seems that the focus in the literature on the formal MPC meeting misses important aspects of the monetary policy process. In particular, the staff seem to play a more important role in forming MPC members’ views than deliberations at the actual committee meeting, and an MPC member’s view of optimal policy seems essentially fixed and known to others prior to the formal meeting. Having established that individuals further down the hierarchy matter for MPC members’ views, we now turn to the role of the one person higher up—the chairman.

5. The Role of the Chairman

The role of the chairman has received considerable attention in the MPC literature.¹⁵ Blinder (2007) distinguishes between two major types of decision-making systems: *individualistic* MPCs and *collegial* MPCs. In individualistic MPCs, the members are individually accountable for their voting, and unanimity is not necessarily sought. The decisions are made by majority voting. Collegial MPCs try to reach unanimity, and if this is not achieved the MPC still stands united behind the decision. There may or may not be a formal vote, and dissenting votes are rare. Blinder distinguishes between two polar types of collegial MPCs: *genuinely collegial* and *autocratically collegial*. The difference relates to the role of the chairman, who merely dictates the “consensus” in an autocratically collegial MPC, while all members have equal weight in a genuinely collegial MPC.

¹⁵Chappell, McGregor, and Vermilyea (2004, 2007a, 2007b) and Gerlach-Kristen and Meade (2010) investigate the influence of the chairman of the FOMC; Chappell, McGregor, and Vermilyea (2014) compare the influence of the chairman in Sweden and the United Kingdom; and Claussen et al. (2012), Gerlach-Kristen (2008), and Riboni and Ruge-Murcia (2010) analyze the role of the chairman in a theoretical setting.

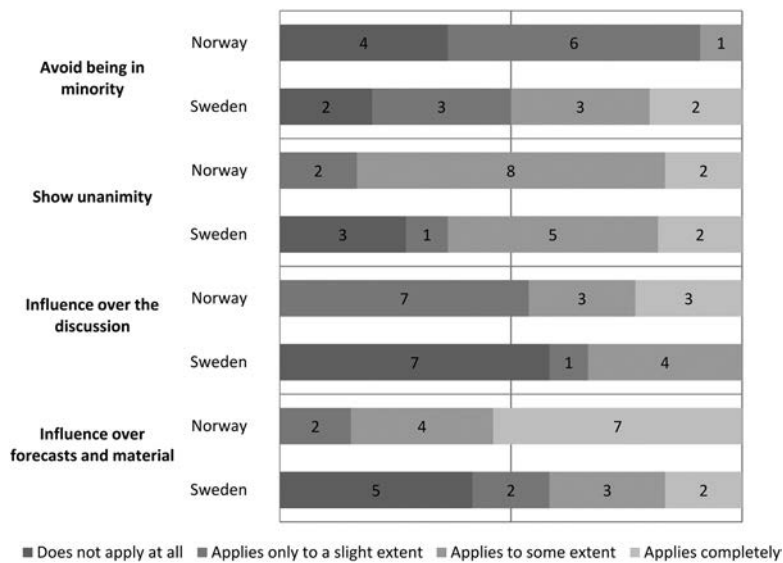
Blinder (2009) cites Norges Bank and the FOMC under Alan Greenspan as examples of autocratically collegial committees. In contrast, the Executive Board of the Riksbank is considered an individualistic committee, as there is formal voting with frequent dissents, and the members are individually accountable for their actions. But even though the Riksbank MPC members are in one sense equal, it is an empirical question whether the chairman is “more equal” than the other members (beyond the fact that he has the casting vote in the case of a tie). It is also an empirical question whether the chairman of the Norges Bank board has a sufficiently dominant role to characterize the board as an autocratically collegial committee.

So, how different are the roles of the chairmen at Norges Bank and the Riksbank? To investigate the role of the chairman, we asked the members to indicate how they thought the following statements, which cover different aspects of chairman dominance, apply:

- *Influence over Forecasts and Material:* The governor has more influence over the forecasts and other materials on which the interest rate decisions are based than the other members of the board.
- *Influence over the Discussion:* As chairman of the monetary policy meeting, the governor is in a particularly good position to influence the discussion, and thereby the interest rate decision.
- *Show Unanimity:* Many members find it worthwhile to show unanimity, and therefore support the governor’s view, given that it is reasonably close to their own assessment.
- *Avoid Being in the Minority:* The governor finds it worthwhile not to be in a minority and therefore supports the majority view, given that it is reasonably close to his own assessment.

We see from figure 9 that most of the Norwegian respondents gave strong support to the statement that the governor has more influence over forecasts and other materials than the other members. They did not, however, agree that the chairman can influence the interest rate decision by steering the discussion. The Riksbank MPC members did not find any of the statements particularly convincing, although the views were somewhat dispersed. The proposition that members

Figure 9. The Role of the Chairman



seek unanimity did, however, receive some support at both central banks, with the median respondent giving it the rating “applies to some extent.”¹⁶

To obtain further evidence of the dominance of the chairman, we asked the members if they have refrained from entering a reservation even if they found that another decision would have been better. Entering a reservation means in this case that a member formally declares in the minutes that he or she does not side with the majority. In a truly individualistic committee, one would expect that members would always enter a reservation if they disagreed with the majority. In a truly collegial committee, one would expect that members form a “consensus” that all can stand behind even if some individuals would prefer a different decision.

Our results showed that seven out of twelve Riksbank MPC members have “never” refrained from entering a reservation if

¹⁶Apel, Claussen, and Lennartsdotter (2010) distinguish between members who were present before and after the publication of the Riksbank’s own interest rate forecasts, and suggest that the introduction of interest rate forecasts has reduced the chairman’s role.

they thought another decision would have been better, while five responded that they have “seldom” refrained. Five out of twelve Norges Bank MPC members responded that they have “never” refrained from entering a reservation, six stated that they have “seldom” refrained, and one responded that he or she has “sometimes” refrained.

Although there is a stronger tendency to refrain from entering a reservation at Norges Bank than at the Riksbank, the difference is not as striking as one might expect. The fact that many Riksbank MPC members have actually refrained from entering a reservation suggests that a committee which from an institutional perspective appears individualistic also has some collegial elements. Similarly, the collegial Norwegian committee seems to have individualistic traits in that most MPC members have never or seldom refrained from entering a reservation.¹⁷ That said, among those members who had “seldom” or “sometimes” refrained from entering a reservation, the argument that won the strongest support was that the final decision was reasonably close to the decision they would have preferred and that there was a “bargaining margin.” Respondents also gave some support to the argument that a reservation was a very strong signal and would hurt the collegial element of the board. One respondent expressed the view that “as an external member, you have a limited ‘budget’ of a few explicit reservations you can make before it comes at the expense of the collegial spirit.”

To summarize this section, our results give ambiguous support to Blinder’s (2009) view that the Norges Bank board is an autocratically collegial committee. Only in one dimension—the governor’s influence over forecasts and material—does the characterization seem to fit well. The Riksbank board seems to fit into the definition of an “individualistic” committee, although with some “collegial” elements. It therefore seems reasonable for theoretical models to take into account the chairman’s influence depending on committee type, but the differences should not be over-emphasized.

¹⁷One consideration in entering a reservation is the long publication lag in Norway. This may, on the one hand, make dissenting attractive in that it does not attract public attention and avoids rifts in the committee; on the other hand, given the low visibility, the lack of public awareness of the reservation means that it cannot be used as a way of lending the dissent more weight.

6. The MPC Members' Preferences (Individual Loss Functions)

The literature not only distinguishes between different kinds of MPCs but also often assumes heterogeneity in MPC members. It is common to assume that MPC members have different individual loss functions; see, e.g., Gerlach-Kristen (2009) and Riboni and Ruge-Murcia (2008).¹⁸ Can MPC members' preferences be captured by a standard quadratic loss function with the inflation gap and the output gap as arguments? Are there significant differences between the members' preferences?

While the existing empirical literature on heterogeneity in MPC members' preferences uses indirect evidence based on voting behavior (e.g., using reaction function estimates; see, e.g., Riboni and Ruge-Murcia 2008), we decided on a direct approach. The reference for the preference questions was the standard loss function

$$L = (\pi - \pi^*)^2 + \lambda(y - y^*)^2, \quad (1)$$

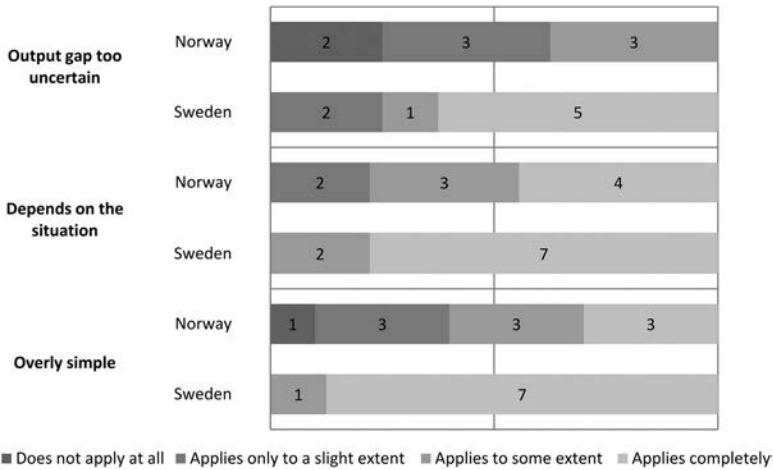
where π is inflation, π^* is the inflation target, $(y - y^*)$ is the output gap (capacity utilization), and λ is the weight on stability in the real economy relative to inflation stability. The respondents were asked if they would be willing to quantify their λ . Only four Riksbank and two Norges Bank MPC members were willing to do this. The typical result was a λ of 1.

Those who responded that they would not consider stating a λ were asked to rate the following three explanations as to why they felt unable to do so:

- *Overly Simple:* The idea that one can use a specific number to describe the relative weight of stabilizing inflation and stabilizing resource utilization is overly simple.
- *Depends on the Situation:* The relative weight depends on the current situation and therefore varies so much that it is meaningless to state a number.

¹⁸MPC members may also differ by their heuristics, i.e., the model they have in mind, as mentioned in section 4.1, or by the economic data available to them.

Figure 10. The MPS Members’ Preferences (individual loss functions)



- *Output Gap Too Uncertain:* The measures of resource utilization in the economy are so uncertain that it is meaningless to state a relative weight.

As seen from figure 10, all of the explanations received considerable support among the Riksbank MPC members, with the median respondent answering “applies completely” for all explanations. Similarly, the median answer for the first and second explanations was “applies to some extent” at Norges Bank. We did not ask whether the quadratic form of the loss function was appropriate, but one of the Norwegian respondents wrote in the questionnaire that “my preferences are not quadratic.” Another member wrote that “I will not give my λ , because it is not always the case that the loss function represents the relevant monetary policy judgments.”

Contrary to the Riksbank MPC members, the Norges Bank MPC members did not give much support to the third explanation, represented by a median answer of “applies to some extent.” This might be due to differences in the extent to which the output gap has been used by the staff and communicated to the MPC. The Norges Bank staff have a long tradition of estimating and publishing their best estimate of the output gap, while the Riksbank staff have been

more agnostic to the output gap as an indicator of capacity during the sample period. The different takes of the two staffs might have influenced the respective MPC members' views on the output gap.

We then asked the members who did not provide a specific figure for λ whether they would be willing to provide a *rough* estimate of how they value stabilizing inflation in relation to stabilizing resource utilization. They were given the following alternatives:

- $\lambda < 1$: I would give *greater* importance to stabilizing inflation than to stabilizing resource utilization.
- $\lambda > 1$: I would give *less* importance to stabilizing inflation than to stabilizing resource utilization.
- $\lambda \approx 1$: I would give roughly *equal* weight to stabilizing inflation and stabilizing resource utilization.

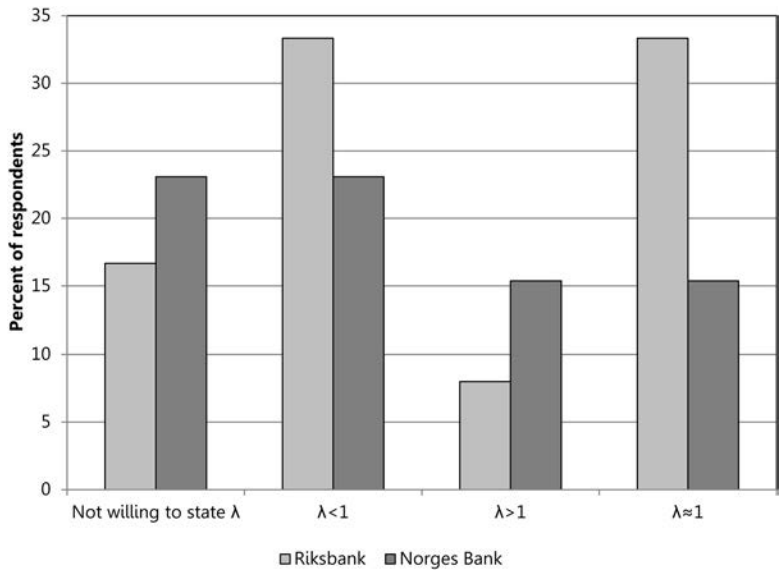
A total of nine Riksbank and eight Norges Bank MPC members were prepared to make either a precise or a rough estimate of the relative importance they would give to stabilizing inflation in relation to stabilizing resource utilization. The results are presented in figure 11. None of the Riksbank MPC members would give greater importance to stabilizing resource utilization than to stabilizing inflation, while two of the Norwegian respondents had such a policy preference. Thus, in our sample, internal members with a full-time position in a central bank are more concerned about price stability than external members with jobs outside the central bank.¹⁹

Perhaps more interesting than the average λ is its dispersion. Preferences seem to be more dispersed at Norges Bank than at the Riksbank. Since members from both central banks have different backgrounds and expertise, this result suggests that heterogeneity is better preserved if members keep their jobs outside the central bank when joining the MPC than if they become full-time internals.²⁰

¹⁹In a similar vein, Gerlach-Kristen (2009) finds that external MPC members at the Bank of England seem to be recession averse, while internal members appear to have symmetric preferences for the output gap. For the United Kingdom, see also Besley, Meads, and Surico (2008) and Spencer (2006); for the United States, see Chappell, McGregor, and Vermilyea (2005).

²⁰In a recent paper related to this issue, Ball (2012) argues that Federal Reserve chairman Ben Bernanke appears to have changed his view about zero-bound policy quite dramatically when becoming a full-time internal at the Federal Reserve.

Figure 11. The Weight on Stabilization of Output Relative to Inflation (λ)



To our knowledge, the theoretical literature does not say whether preference dispersion is good or bad. On the one hand, from a democratic accountability perspective, it may be desirable that the MPC roughly reflects the preference distribution in society as a whole. On the other hand, there are arguments for having MPC members with preferences that deviate from those of society in certain ways, cf. Rogoff’s (1985) “conservative” central banker. Notwithstanding, our results indicate that the choice of external vs. internal MPC members may have implications for the degree of preference dispersion within the MPC.²¹

In sum, this section suggests considerable heterogeneity in terms of preferences across MPC members at both the Riksbank and Norges Bank. Theoretical papers assuming such differences seem to be on the right track.

²¹One concern with preference dispersion is that it may be more difficult to reach a decision to change interest rates, particularly if decisions are made unanimously (see Gerlach-Kristen 2005). The problem is reduced in committees that use other decision mechanisms.

7. Transparency

Transparency is one final field in the institutional setup of MPCs that our survey can shed light on. There is a large general literature on monetary policy transparency.²² The theoretical results on the benefits of transparency are ambiguous, but a rough general conclusion is that neither full secrecy nor full transparency is optimal. We first discuss transparency in terms of public speeches and then in terms of the publication of meeting minutes.

7.1 *Speeches*

Where monetary policy is set by a single governor, it is clear that his speeches carry the most weight of all the verbal central bank communications. But who should speak to the public in an MPC, and how, if at all, should speeches be coordinated? Blinder (2007, p. 114) writes: “A central bank that speaks with a cacophony of voices may, in effect, have no voice at all.”

To learn about how the two groups of MPC members view the relevance of the “cacophony problem,” we asked them if they thought the following statements about individual communication applied:

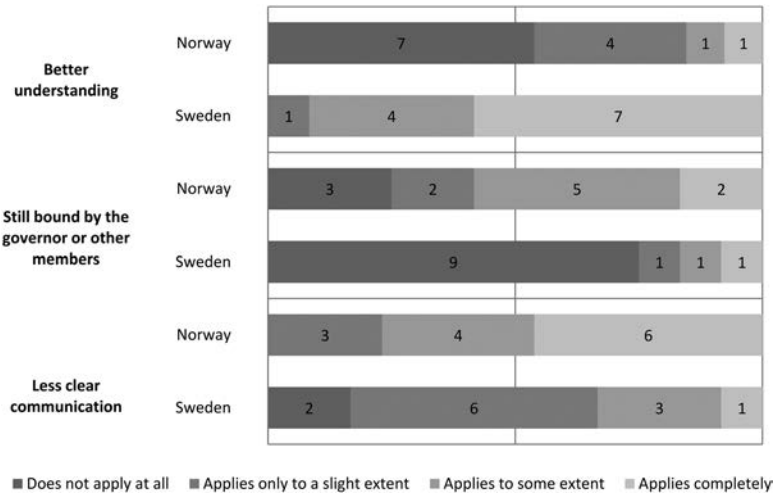
- *Less Clear Communication:* If all members of the Executive Board can express their own views publicly as individuals, the Bank’s communication is less clear.
- *Still Bound by the Governor or Other Members:* When expressing my own views publicly, I feel bound by the governor’s or other members’ statements.²³
- *Better Understanding:* When each member describes his or her views in public (e.g., in speeches and the minutes of meetings), agents better understand the bank’s policy (“the way we think”).

The responses are represented in figure 12. There are considerable differences between the two groups of policymakers, basically

²²For reviews of the literature on transparency, see, e.g., Blinder et al. (2008), Crowe and Meade (2008), and Geraats (2002, 2006). See also Chahrour (2011).

²³The exact wording was “the governor’s/deputy governor’s statements” for Norges Bank and “other members’ statements” for the Riksbank.

Figure 12. One Voice vs. Many Voices



following the collegial vs. individualistic setup of the two committees. The Norges Bank MPC members tended to believe in Blinder’s “cacophony problem,” while the Swedish respondents thought publishing the different views improves agents’ understanding. The contrast was particularly stark regarding the question of whether such information makes economic agents understand the bank’s policy better, which was a view strongly supported in Sweden. There was also a difference in how the members would feel bound by the governor’s or other members’ statements, where the median Norges Bank MPC member felt more constrained than his or her counterpart at the Riksbank.²⁴

There is thus a strong tendency to defend the respondent’s own system, i.e., a status quo bias. Apparently, members of both MPCs think that their way of communicating works satisfactorily and need not be adjusted. This implicitly suggests the exact communication strategy is not crucial, which would support the finding in Ehrmann and Fratzscher (2007) that the predictability of policy decisions is

²⁴We should add, however, that there was considerable disagreement within each committee on this issue, where the responses varied from “does not apply at all” to “applies completely” in each group.

equally good for the Federal Reserve and the ECB, even though these banks have chosen an individualistic and a collegial communication strategy, respectively.

7.2 *Minutes*

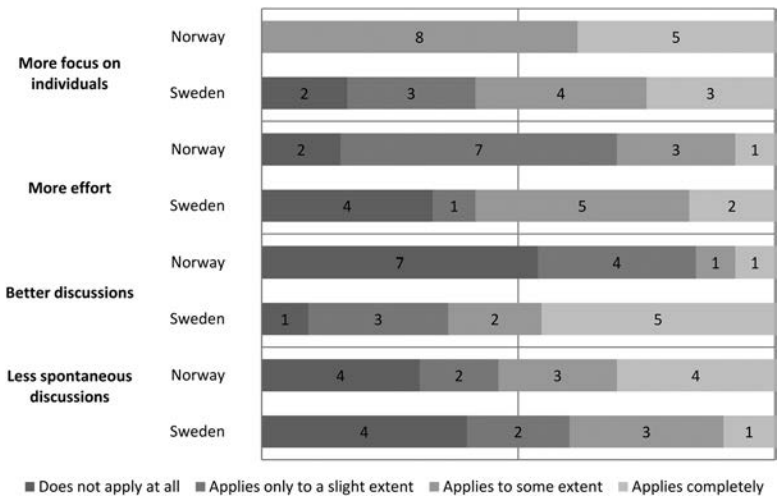
Whether or not minutes and the voting results of the monetary policy meeting should be published is a controversial issue. The pros and cons of revealing information about the deliberations of individual MPC members—for example, by publishing individual voting records—have been explored in a number of theoretical papers (see, for instance, Gersbach and Hahn 2008, 2009, and Weber 2010).

Empirical results suggest that publishing voting records makes it easier to predict future policy decisions (see, e.g., Gerlach-Kristen 2004, 2009). Also, publishing voting records and attributed views could give incentives to prepare better for meetings, particularly if this increases the probability of reappointment.

However, as with verbal communication, the concern with written communication is that publishing minutes and voting results may just confuse matters. Issing (2005), for example, fears that economic agents will attach more importance to MPC members' individual opinions than to the relevant economic arguments. A second argument against publishing voting results and minutes—in particular, attributed minutes—is that it might have the effect of making the discussion at the policy meeting less honest and frank, and more limited and “tied to script.” Meade and Stasavage (2008) provide some evidence from the Federal Reserve's Federal Open Market Committee that publishing verbatim transcripts of the FOMC meetings made members more reluctant to offer dissenting opinions.

Yet another possible consequence of publishing minutes that has been explored in one strand of the literature is that MPC members worry that the public might interpret disagreement as a sign of incompetence. In this case, committee members may want to meet to discuss—and possibly settle—things before the monetary policy meeting, away from the public eye (see, e.g., Swank, Swank, and Visser 2008 and Swank and Visser 2008).

Figure 13. Attributed Minutes



To investigate how actual MPC members view some of the possible consequences of publishing this information for the quality of the deliberations, we asked them to rate the following statements:

- *Less Spontaneous Discussions:* The discussion at monetary policy meetings is probably less spontaneous or in some other way poorer when the minutes are attributed.
- *Better Discussions:* The discussions at monetary policy meetings are probably better when the minutes are attributed.
- *More Effort:* The members probably invest more time and effort in monetary policy preparation work when the minutes are attributed.
- *More Focus on Individuals:* There is probably more public focus on the individual members when minutes are attributed.

Figure 13 shows the responses. Although there was substantial disagreement within each committee, there was a tendency for members to defend the status quo. That is, members from the Riksbank, which has attributed minutes, were sympathetic to the argument that attributed minutes make the discussions better and that members invest more time and effort. Members from Norges Bank, which

does not have attributed minutes, did not think that discussions would be better if they were provided. At the same time, they acknowledged that attributed minutes would make members prepare better.

In sum, the survey respondents' views on how much transparency is desirable are mixed. As regards verbal communication, the members of the individualistic MPC at the Riksbank think that if all members express their views, this helps the public understand monetary policy better, whereas the collegial MPC members in Norway do not hold this view.²⁵ Nor is there a common view on whether publishing minutes would be beneficial. Generally, members preferred the status quo in their respective committee.

8. Discussion and Conclusions

The literature on MPCs is continually expanding. In this paper we contribute to this literature by using a survey method to obtain information on how MPCs work in practice.

We find that many assumptions and results in the previous literature are confirmed by actual MPC members' experience: a committee is preferable to a single policymaker but it should not be very large, MPC members do indeed have heterogeneous preferences, and the chairman seems to have more influence in a collegial than in an individualistic committee.

However, there are also results that are less intuitive and quite at odds with the conventional, stylized view of monetary policymaking by committees. One striking finding is that there seems to be only limited pooling within the committee, in the sense that members do not seem to get much input from, nor are they influenced by, each other. This result was found in the main survey where we asked about pooling of three types of information, and confirmed in the follow-up study where we asked about pooling more generally.

²⁵As suggested by a referee, this is an issue where there may be a certain non-response bias, as the four (out of seventeen) members who did not respond at Norges Bank may be less sympathetic to transparency than the average member (as they refrained from responding). If so, the aversion to transparency is somewhat under-estimated. Unfortunately, we have not been able to investigate this further.

We draw two conclusions from this. One is that the benefits of committee decision making that members do report are more due to “pooling by voting” than to “pooling by talking.” The other is that the limited amount of pooling is likely to be related to the institutional features and traditions of the two banks. In neither bank do MPC members have their own dedicated staff. The work on collecting new information, analyzing this information, doing research, building models, producing a majority forecast, etc. is conducted by the bank’s staff. Thus, there might simply not be much scope for pooling left. Indeed, the crucial role of the bank’s staff was confirmed by the respondents. It would be an interesting issue for future research to see if there is more pooling among MPC members in committees with individual staff and/or no common forecasts.

Another finding is that there is a preference for the status quo in several areas where the institutional setups of the two banks differ—for example, the choice of communication strategy and whether or not voting records should be published. This does not come as a surprise, as it is largely up to the committees themselves to decide how to communicate and how transparent they should be. These are also issues where the literature is inconclusive.

The two MPCs also had different views on whether or not the committee should include external members. The Swedish respondents tend to favor a fully internal MPC whereas the Norwegian respondents advocate an MPC including external members. Even if it cannot be ruled out that the Riksbank’s members under-estimate the risks of groupthink, a more plausible explanation for the diverging results is once again that they are affected by the practices and prevailing institutional frameworks in the two banks. The Riksbank’s MPC members have with few exceptions been recruited from outside the bank, typically from the financial sector or academia (the same holds for Norges Bank). In addition, many members spend only one term (six years) on the MPC. Thus, they might keep much of the outsider’s perspective while serving on the MPC. Moreover, the Riksbank has quite high ambitions regarding transparency, manifested in the publication of attributed minutes and voting records. A high degree of transparency, where individual MPC members’ arguments and reasoning are continuously exposed to public scrutiny, is a further safeguard against groupthink. In this type of setup, there may be little to gain by bringing external members onto the MPC.

At Norges Bank, the setting is different with collegial accountability and less openness. It could be argued that the lesser degree of openness makes external members more important. In a sense, external members may fulfill the same function as transparency—as a check and balance against an MPC getting carried away, because of groupthink or for other reasons. Thus, considering the differences in the respective frameworks, the different views on external MPC members in Norway and Sweden do not appear very surprising either.

The tendency to support the status quo in their respective committee indicates, of course, that members are content with the present institutional setup and think it works satisfactorily. This, in turn, implicitly suggests that there is no one obviously preferable choice in these matters.²⁶

To conclude, the pivotal idea in this paper is to get information on monetary policy decision making “straight from the horse’s mouth” by asking present and former MPC members about their experiences and views. The kind of information yielded would have been difficult to obtain in other ways and provides insights into the inner workings and thinking of MPCs. Although surveys have possible shortcomings, we believe that the results in this study add to the knowledge about decision making in monetary policy committees and we hope that they will inform future research.

References

- Amirault, D., C. Kwan, and G. Wilkinson. 2004. “A Survey of the Price-Setting Behaviour of Canadian Companies.” *Bank of Canada Review* (Winter 2004–2005): 29–40.
- Apel, M., C. A. Claussen, and P. Lennartsdotter. 2010. “Picking the Brains of MPC Members.” Working Paper No. 237, Sveriges Riksbank.

²⁶If the Swedish MPC members differ systematically from their Norwegian counterparts, it could be wrong to attribute different survey answers in the two countries to the MPC design or institutional differences in the two countries. However, we have no reason to believe that there are systematic differences in the characteristics or personalities of the members of the two groups.

- Apel, M., R. Friberg, and K. Hallsten. 2005. "Microfoundations of Macroeconomic Price Adjustment: Survey Evidence from Swedish Firms." *Journal of Money, Credit and Banking* 37 (2): 313–38.
- Ball, L. M. 2012. "Ben Bernanke and the Zero Bound." NBER Working Paper No. 17836.
- Berger, H., and V. Nitsch. 2008. "Too Many Cooks? Committees in Monetary Policy." Working Paper No. 2274, CESifo Group Munich.
- Berk, J. M., and B. K. Bierut. 2011. "Communication in a Monetary Policy Committee." *European Journal of Political Economy* 27 (4): 791–801.
- Besley, T., N. Meads, and P. Surico. 2008. "Insiders versus Outsiders in Monetary Policymaking." *American Economic Review: Papers and Proceedings* 98 (2): 218–23.
- Blinder, A. S. 2000. "Central-Bank Credibility: Why Do We Care? How Do We Build It?" *American Economic Review* 90 (5): 1421–31.
- . 2006. "Monetary Policy Today: Sixteen Questions and about Twelve Answers." Paper prepared for the Banco de España conference on "Central Banks in the 21st Century," Madrid, June 8–9.
- . 2007. "Monetary Policy by Committee: Why and How?" *European Journal of Political Economy* 23 (1): 106–23.
- . 2009. "Talking about Monetary Policy: The Virtues (and Vice?) of Central Bank Communication." BIS Working Paper No. 274.
- Blinder, A. S., E. R. D. Canetti, D. E. Lebow, and J. B. Rudd. 1998. *Asking About Prices — A New Approach to Understanding Price Stickiness*. New York: Russell Sage Foundation.
- Blinder, A. S., M. Ehrmann, M. Fratzscher, J. De Haan, and D.-J. Jansen. 2008. "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence." *Journal of Economic Literature* 46 (4): 910–45.
- Blinder, A. S., and J. Morgan. 2005. "Are Two Heads Better Than One? Monetary Policy by Committee." *Journal of Money, Credit and Banking* 37 (5): 789–812.
- . 2008a. "Do Monetary Policy Committees Need Leaders? A Report on an Experiment." *American Economic Review* 98 (2): 224–29.

- . 2008b. "Leadership in Groups: A Monetary Policy Experiment." *International Journal of Central Banking* 4 (4): 117–50.
- Boulier, B., and R. Goldfarb. 1998. "On the Use and Nonuse of Surveys in Economics." *Journal of Economic Methodology* 5 (1): 1–21.
- Chahrour, R. 2011. "Public Communication and Information Acquisition." Manuscript, Columbia University.
- Chappell, H. W. Jr., R. R. McGregor, and T. Vermilyea. 2004. "Majority Rule, Consensus Building, and the Power of the Chairman: Arthur Burns and the FOMC." *Journal of Money, Credit and Banking* 36 (3): 407–22.
- . 2005. *Committee Decisions on Monetary Policy: Evidence from Historical Records of the Federal Open Market Committee*. MIT Press.
- . 2007a. "The Persuasive Power of a Committee Chairman: Arthur Burns and the FOMC." *Public Choice* 132 (1–2): 103–12.
- . 2007b. "The Role of the Bias in Crafting Consensus: FOMC Decision Making in the Greenspan Era." *International Journal of Central Banking* 3 (2): 39–60.
- . 2012. "Deliberation and Learning in Monetary Policy Committees." *Economic Inquiry* 50 (3): 839–47.
- . 2014. "Power-Sharing in Monetary Policy Committees: Evidence from the United Kingdom and Sweden." *Journal of Money, Credit and Banking* 46 (4): 665–92.
- Claussen, C. A., E. Matsen, Ø. Røisland, and R. Torvik. 2012. "Overconfidence, Monetary Policy Committees and Chairman Dominance." *Journal of Economic Behavior and Organization* 81 (2): 699–711.
- Crowe, C., and E. E. Meade. 2008. "Central Bank Independence and Transparency: Evolution and Effectiveness." *European Journal of Political Economy* 24 (4): 763–77.
- Dincer, N., and B. Eichengreen. 2013. "Central Bank Transparency and Independence: Updates and New Measures." Working Paper No. 2013-21 (2013.09), Bank of Korea.
- 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.
- Eijffinger, S. C. W., and P. M. Geraats. 2006. "How Transparent are Central Banks?" *European Journal of Political Economy* 22 (1): 1–21.

- Erhart, S., and J.-L. Vasquez-Paz. 2007. "Optimal Monetary Policy Committee Size: Theory and Cross Country Evidence." Working Paper No. 2007/6, Central Bank of Hungary.
- Fabiani, S., M. Druant, I. Hernando, C. Kwapił, B. Landau, C. Loupias, F. Martins, T. Mathä, R. Sabbatini, H. Stahl, and A. Stokman. 2006. "What Firms' Surveys Tell Us about Price-Setting Behavior in the Euro Area." *International Journal of Central Banking* 2 (3): 3–47.
- Geraats, P. 2002. "Central Bank Transparency." *Economic Journal* 112 (483): 532–65.
- . 2006. "Transparency of Monetary Policy: Theory and Practice." *CESifo Economic Studies* 52 (1): 111–52.
- Gerlach-Kristen, P. 2004. "Is the MPC's Voting Record Informative about Future UK Monetary Policy?" *Scandinavian Journal of Economics* 106 (2): 299–313.
- . 2005. "Too Little, Too Late: Interest Rate Setting and the Costs of Consensus." *Economics Letters* 88 (3): 376–81.
- . 2006. "Monetary Policy Committees and Interest Rate Setting." *European Economic Review* 50 (2): 487–507.
- . 2008. "The Role of the Chairman in Setting Monetary Policy: Individualistic vs. Autocratically Collegial MPCs." *International Journal of Central Banking* 4 (3): 119–43.
- . 2009. "Outsiders at the Bank of England's MPC." *Journal of Money, Credit and Banking* 41 (6): 1099–1115.
- Gerlach-Kristen, P., and E. Meade. 2010. "Is There a Limit on FOMC Dissents? Evidence from the Greenspan Era." Working Paper No. 2010-16, American University, Department of Economics.
- Gerling, K., H. P. Grüner, A. Kiel, and E. Schulte. 2005. "Information Acquisition and Decision Making in Committees: A Survey." *European Journal of Political Economy* 21 (3): 563–97.
- Gersbach, H., and V. Hahn. 2008. "Should the Individual Voting Records of Central Bankers be Published?" *Social Choice and Welfare* 30 (4): 655–83.
- . 2009. "Voting Transparency in a Monetary Union." *Journal of Money, Credit and Banking* 41 (5): 831–53.
- Hall, S., M. Walsh, and A. Yates. 2000. "Are UK Companies' Prices Sticky?" *Oxford Economic Papers* 52 (3): 425–46.

- Hallsten, K., and S. Tägtström. 2009. "The Decision-Making Process — How the Executive Board of the Riksbank Decides on the Repo Rate." *Sveriges Riksbank Economic Review* 2009 (1): 69–86.
- Issing, O. 2005. "Communication, Transparency, Accountability: Monetary Policy in the Twenty-First Century." *Review* (Federal Reserve Bank of St. Louis) 87 (2, Part 1): 65–83.
- Langbraaten, N., E. W. Nordbø, and F. Wulfsberg. 2008. "Price-Setting Behaviour of Norwegian Firms — Results of a Survey." *Economic Bulletin* (Norges Bank) 2/2008: 13–34.
- Lombardelli, C., J. Proudman, and J. Talbot. 2005. "Committees Versus Individuals: An Experimental Analysis of Monetary Policy Decision-Making." *International Journal of Central Banking* 1 (1): 181–205.
- Meade, E. E., and D. Stasavage. 2008. "Publicity of Debate and the Incentive to Dissent: Evidence from the US Federal Reserve." *Economic Journal* 118 (528): 695–717.
- Riboni, A., and F. J. Ruge-Murcia. 2008. "Preference Heterogeneity in Monetary Policy Committees." *International Journal of Central Banking* 4 (1): 213–33.
- . 2010. "Monetary Policy by Committee: Consensus, Chairman Dominance, or Simple Majority?" *Quarterly Journal of Economics* 125 (1): 363–413.
- Rogoff, K. 1985. "The Optimal Degree of Commitment to an Intermediate Target." *Quarterly Journal of Economics* 100 (4): 1169–90.
- Romer, C. D., and D. H. Romer. 2008. "The FOMC versus the Staff: Where Can Monetary Policymakers Add Value?" *American Economic Review* 98 (2): 230–35.
- Sibert, A. 2006. "Central Banking by Committee." *International Finance* 9 (2): 145–68.
- Spencer, C. 2006. "Reaction Functions of Bank of England MPC Members: Insiders versus Outsiders." Working Paper No. 06-06, University of Surrey, School of Economics.
- Svensson, L. E. O. 2009. "Transparency under Inflation Targeting: Experiences and Challenges." *Sveriges Riksbank Economic Review* 2009 (1): 5–44.
- Swank, J., O. H. Swank, and B. Visser. 2008. "How Committees of Experts Interact with the Outside World: Some Theory, and

Evidence from the FOMC.” *Journal of the European Economic Association* 6 (2–3): 478–86.

Swank, O. H., and B. Visser. 2008. “Is Transparency to No Avail? Committee Decision-Making, Pre-Meeting, and Credible Deals.” Working Paper No. ECO 2008/18, European University Institute.

Weber, A. 2010. “Communication, Decision Making, and the Optimal Degree of Transparency of Monetary Policy Committees.” *International Journal of Central Banking* 6 (3): 1–49.

The ECB Unconventional Monetary Policies: Have They Lowered Market Borrowing Costs for Banks and Governments?*

Urszula Szczerbowicz
CEPII

This paper evaluates the impact of the ECB's unconventional policies on bank and government borrowing costs. We employ event-based regressions to assess and compare the effects of asset purchases and exceptional liquidity announcements on money markets, covered bond markets, and sovereign bond markets. The results show that (i) exceptional liquidity measures (three-year loans to banks and setting the ECB deposit rate to zero) significantly reduced persistent money-market tensions and that (ii) asset purchases were the most effective in lowering refinancing costs of banks and governments in the presence of high sovereign risk. Moreover, we show that the ECB asset purchases fed through into other asset prices: bank covered bond purchases diminished sovereign spreads, while sovereign bond purchases reduced covered bond spreads.

JEL Codes: E43, E44, E52, E58.

*I thank Pierpaolo Benigno, Nicola Borri, Gunther Capelle-Blancard, Virginie Coudert, Kaku Furuya, Erwan Gautier, Thomas Grjebine, Francesco Giavazzi, Refet Gürkaynak, Olena Havrylchyk, Minoru Kaneko, Atsushi Nakajima, Etsuro Shioji, Henri Sterdyniak, Michio Suzuki, Willem Thorbecke, Fabien Tripier, Taisuke Uchino, and Kazuo Ueda for their valuable advice. I am also grateful for helpful comments and suggestions to participants at the CEPR 22nd ESSIM conference, RIETI Discussion Paper seminar, Tokyo University macro seminar, Daito Bunka University seminar, T2M 2013, 62nd annual meeting of the French Economic Association, 28th annual congress of the European Economic Association, and the CEPII/PSE Macro Finance Workshop. The paper was finalized during the research stay at the RIETI, whose excellent working environment is greatly appreciated. Author contact: CEPII, 113 rue de Grenelle, 75007 Paris, France. Tel: +33 (0)1 53 68 55 35. Fax: +33 (0)1 53 68 55 01. E-mail: urszula.szczerbowicz@cepii.fr.

1. Introduction

Following the 2007–9 financial crisis, the euro zone was further hit by the sovereign debt crisis that started in Greece and spread to other member countries. The debt crisis led to the fragmentation of the single financial market and resulted in important differences in credit conditions across the euro-zone states. The situation was further deepened by the negative feedback loop between the sovereign distress and bank insolvency. Indeed, euro-zone banks were heavily exposed to sovereign debt, while euro-zone governments bore the responsibility of rescuing their banking systems. The European Central Bank (ECB) faced the difficult task of restoring monetary transmission to support the economy in these exceptional circumstances. However, the traditional monetary tool—the ECB main refinancing rate—was not effective in equalizing the borrowing conditions across the euro zone and stabilizing the malfunctioning interbank market. Therefore, the ECB implemented several unconventional monetary policies to attain its goals.

The objective of this paper is to evaluate the effectiveness of these unconventional monetary policies. We employ event-based regressions to measure the impact of the ECB policies on the market borrowing costs for banks and governments in the euro-zone countries. To this purpose, we create a timeline of unconventional monetary policy announcements and classify them into distinct categories: long-term sovereign bond purchases (Securities Markets Programme, SMP), short-term sovereign bond purchases (Outright Monetary Transactions, OMT), covered bond purchase programs (CBPP1 and CBPP2), three-year long-term refinancing operations (three-year LTROs), lowering the ECB deposit rate to 0 percent, and unlimited liquidity provisions (the fixed-rate full-allotment procedure, FRFA). By using the Factiva press database, we make sure to include other important events that occurred on monetary policy announcement days. Finally, we test the impact of all announcements on money-market spreads, covered bond spreads, and sovereign bond spreads.

We find that three-year LTROs and cutting the ECB deposit rate to 0 percent significantly reduced persistent money-market tensions in the euro zone. Furthermore, our results show that central bank interventions in sovereign markets (SMP and OMT) are particularly effective when sovereign risk is high: periphery euro-zone

countries benefited the most from the sovereign bond purchasing programs, while the French spread reacted very little if at all. The strong impact in the troubled periphery countries confirms the ECB's capacity to establish more homogeneous credit conditions in the euro zone. Interestingly, the effects of the ECB's asset purchases were not limited to the market of intervention. *Sovereign* bond purchases impacted not only *sovereign* bond spreads but also bank *covered* bond spreads. Likewise, *covered* bond purchases diminished not only *covered* bond spreads but also *sovereign* bond spreads. The presence of a feedback loop between sovereigns and the banking sector in the euro zone might be a source of this additional effect of the ECB asset purchases.

There exists an important literature about the effects of unconventional monetary policies in the United States, the United Kingdom, and Japan.¹ The empirical evidence of the impact of the ECB's unconventional monetary policies on financial markets is also growing rapidly.² In the following literature review, we focus on the three markets relevant for our analysis—money markets, covered bond markets, and sovereign bond markets—and explain how we contribute to the literature.

The impact of the ECB exceptional liquidity measures on inter-bank lending was studied via regression analysis by Abbassi and Linzert (2011), Angelini, Nobili, and Picillo (2011), and Brunetti, di Filippo, and Harris (2011), who did not find significant effects of exceptional refinancing operations up to one year on money-market tensions.³ We complement these papers by showing that only

¹For the United States see, for instance, Hamilton and Wu (2012), Krishnamurthy and Vissing-Jorgensen (2011), Szczerbowicz (2011), or Taylor and Williams (2009); for the United Kingdom, see Joyce et al. (2011); and for Japan, see Ueda (2012, 2013).

²We present here the impact of the ECB unconventional monetary policies on financial markets, but there exists also a large literature on their effects on macroeconomic variables (see, for instance, Lenza, Reichlin, and Pill 2010 or Peersman 2011).

³Abbassi and Linzert (2011) find that three-month special LTROs diminished the three-month and six-month Euribor (significance at 10 percent), twelve-month special LTROs diminished only the twelve-month Euribor (significance at 5 percent), and six-month LTROs were not significant at all. However, when the sample is split into two periods (August 2007–October 2008 and October 2008–June 2009), none of the special LTROs is significant. Angelini, Nobili, and Picillo

stronger liquidity measures (three-year LTROs and 0 percent deposit rate at the ECB) reduced significantly interbank distress.

Concerning covered bond markets, Beirne et al. (2011) evaluated via an event study and regression analysis the impact of the first covered bond purchasing program (CBPP1) and found that it was effective in lowering covered bond spreads. Our paper confirms this effect not only for CBPP1 but also for the second covered bond purchasing program (CBPP2). Moreover, we show the effectiveness of the covered bond purchases in reducing *sovereign* bond market distress.

There are also several studies that measure the impact of the first sovereign bond purchasing program (SMP) on sovereign bond markets. In this paper, we find that the SMP announcement reduced greatly the spreads in the periphery euro-zone countries. Independently, Eser and Schwaab (2013), Ghysels et al. (2013), and Pattipeilohy et al. (2013) find similar announcement results. Moreover, De Pooter, Martin, and Pruitt (2012), Eser and Schwaab (2013), and Ghysels et al. (2013) show that the actual SMP operations were also effective in reducing sovereign spreads. We add to this literature by providing new evidence on the second sovereign bond purchase program (OMT), which proved effective in reducing sovereign market tensions in the peripheral euro zone.⁴ Interestingly, we also find that both sovereign bond purchase programs (OMT and SMP) were very effective in reducing *covered* bond market distress and therefore lowering longer-term bank funding costs.

Finally, we show that simultaneous analysis of asset purchases and liquidity measures on bank and government borrowing costs is essential for a better understanding of the unconventional monetary policies' effectiveness in the presence of sovereign risk. In particular, by studying both bank and sovereign funding markets, we are

(2011) find that one-month and three-month special LTROs diminish the long-term interbank spread only after the failure of Lehman Brothers (three-month LTROs are significant at 5 percent, while one-month LTROs are significant at 10 percent), and six-month LTROs had an unexpected positive sign and increased the interbank spread. Brunetti, di Filippo, and Harris (2011) show further that special three-month LTROs increased both the level of the bid-ask spread and its volatility, contributing to higher uncertainty in money markets.

⁴Recently, Falagiarda and Reitz (2013) measured the announcement effect for the OMT on Italian spreads.

able to capture the significant impact of the ECB sovereign asset purchases on bank borrowing cost, which might be explained by the sovereign-banking nexus in the euro zone.

This finding has important policy implications. First, the ECB can reduce the sovereign spreads not only through sovereign bond purchases but also through bank covered bond purchases, which has important implications given the political and operational constraints of the ECB. Second, it can diminish bank long-term refinancing costs and lower covered bond spreads by purchasing either covered or sovereign bonds. Sovereign bond purchases are in fact even more effective, suggesting that the increase in bank covered bond spreads was mostly due to increased sovereign default probability.

The remainder of this paper is organized as follows. The ECB's unconventional monetary policy announcements and their objectives are described in section 2. Methodology and data are presented in section 3. In section 4 we estimate the impact of the ECB's announcements on money-market, covered bond, and sovereign bond spreads. Section 5 concludes.

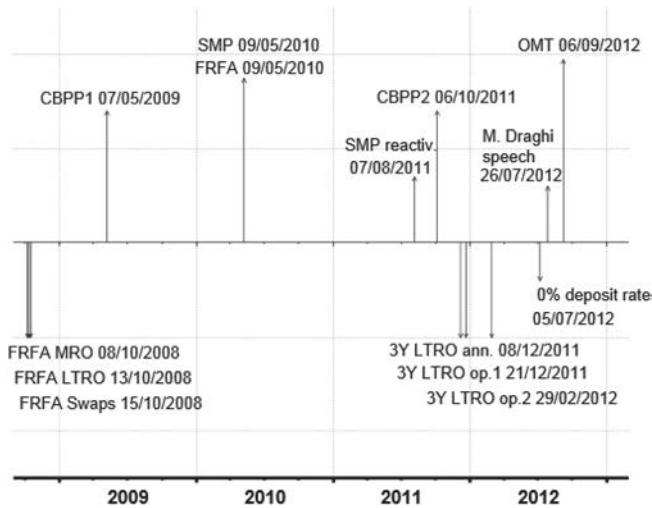
2. The ECB's Unconventional Monetary Policies

This section presents the ECB's unconventional policies, their theoretical foundations, and the objectives they were meant to attain. Even though we do not test the transmission channels of these policies in the empirical section, the theoretical background is important for the interpretation of the results. We regroup unconventional policies into two categories: (i) exceptional liquidity provisions (three-year LTROs, the fixed-rate full-allotment procedure, and setting the deposit rate to zero) and (ii) asset purchases (sovereign bond and covered bond purchase programs). Figure 1 presents a timeline of non-conventional monetary policy announcements considered in this paper.

2.1 Exceptional Liquidity Provisions

Significant tensions appeared on the euro-zone interbank market at the onset of the subprime crisis. The general uncertainty concerning banks' balance sheet health led to the increase in the spread between the risky interbank rate (Euribor) and the riskless rate

Figure 1. Timeline of the Unconventional Monetary Policy Announcements



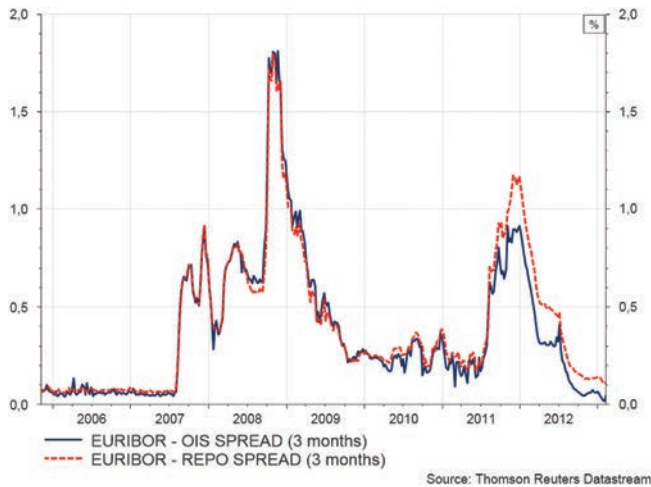
Notes: All the announcements are described in table 5 in the appendix. Dates shown in this figure are in day/month/year format.

(figure 2). The euro-zone sovereign debt crisis further impaired the money-market functioning, as the banks held important amounts of risky sovereign debt issued by periphery euro-zone countries.

The ECB reacted very promptly to the tensions on the interbank market and implemented several additional liquidity measures. In this paper, we focus on the impact of the strongest ECB liquidity innovations: announcements of the fixed-rate full-allotment procedure (FRFA) and the three-year refinancing operations (three-year LTROs). We also consider the announcement of setting the ECB deposit rate to zero, as it was the first time the ECB hit this limit.

The fixed-rate procedure with full allotment was also a part of the ECB's non-standard toolbox. Traditionally, open-market operations were conducted through variable-rate tenders.⁵ Under the FRFA

⁵It should be noted, however, that the fixed tenders for the main refinancing operations (MROs) were applied in the beginning of the Eurosystem (from January 1999 to June 2000) and are compatible with the conventional monetary policy framework (Ayuso and Repullo 2003).

Figure 2. Money-Market Spreads in the Euro Zone

procedure, banks could satisfy all their liquidity needs at an interest rate specified in advance (the interest rate on the main refinancing operations). After Lehman Brothers collapsed, the ECB introduced the FRFA procedure for all open-market operations and for foreign liquidity swaps. First, late on October 8, 2008, the ECB announced that all weekly main refinancing operations (MROs) would be carried out through a fixed-rate tender procedure with full allotment. On October 13, 2008, it decided to provide unlimited dollar funding in coordinated action with the Federal Reserve. Two days later, on October 15, 2008, the ECB announced an FRFA procedure for its LTROs. The ECB decided to return to a variable-rate tender procedure in the regular three-month LTROs in March 2010, but the Greek debt crisis forced it to resume the FRFA procedure in the regular LTROs on May 10, 2010. By ensuring banks' continued access to liquidity, the ECB intended to offset liquidity risk in the market.

Since 2007 the ECB has implemented other exceptional liquidity measures: gradual lengthening of the maturity of the LTROs up to one year. These liquidity provisions are very close to standard monetary measures and were often expected by the market participants. However, on December 8, 2011, the ECB took an unprecedented measure to conduct three-year LTROs as a fixed-rate procedure with full allotment. The first three-year LTRO was offered on December

21, 2011, and the second on February 29, 2012. The banks borrowed more than €1 trillion, which covered their immediate funding needs and prevented them from selling assets and curtailing some types of lending. The three-year LTROs were incomparable in length to other liquidity measures and considerably increased the credit risk on the ECB balance sheet.

The main objective of the ECB exceptional liquidity provisions was to restore the smooth functioning of the interbank market, as this aspect was crucial for extending credit to firms and households. The liquidity measures can be effective in stabilizing the interbank market for several reasons. A liquidity shortage has a negative impact on financial institution lending capabilities and may result in a credit crunch. Liquidity-constrained banks excessively hoard liquidity for precautionary reasons and proceed to fire sales of assets, affecting negatively their prices. By ensuring funding liquidity, the ECB's unconventional measures diminish these adverse effects. They also reduce banks' uncertainty with respect to funding liquidity of other market participants and therefore diminish counterparty risk premiums.

Despite unlimited liquidity being available, the interbank market was still not functioning. In order to overcome banks' reluctance to lend to each other, the ECB lowered its deposit rate to 0 percent on July 5, 2012. While the markets expected a cut in the deposit rate on that day, the move to zero was a surprise.⁶ This measure was not a strictly unconventional measure, but it was the first time that the ECB hit the zero bound, and it was perceived as moving into "new territory." While not a liquidity measure *per se*, it was aimed at reinforcing the existing liquidity tools by encouraging banks to lend available money in the interbank market and not store it at the ECB.

2.2 Purchases of Assets

In a period of financial distress, the central bank can modify the composition of its assets by purchasing the securities that suffer

⁶From "Euro Hurt by ECB Rate Cuts," *Financial Times*, July 6, 2012: "The euro fell sharply to its lowest level . . . after the European Central Bank cut its main interest rate as expected to 0.75 per cent and, in a surprise move, the rate on its deposit facility to zero." See also "Euribor Rates Fall Ahead of ECB Rate Decision," Reuters News, July 5, 2012.

from temporary liquidity problems or are undervalued by financial markets. This policy is sometimes called “credit easing.” The purchases can be sterilized by disposal of the other central bank assets (“pure credit easing”) or be a part of the central bank balance sheet expansion (“quantitative easing”).

The effectiveness of credit easing is based on the “portfolio rebalancing effect”: when securities are not perfect substitutes, reducing the quantity of selected assets available for private investors increases their prices and diminishes yields by suppressing the risk premia (Bernanke 2010). The portfolio rebalancing effect is controversial from a theoretical point of view. A representative-agent model of Eggertsson and Woodford (2003) predicts no effect for such operations on price level or output. However, replacing a representative agent that has no preference between markets and assets with heterogeneous agents can also provide rationale for central bank asset purchasing. In the preferred-habitats model of Vayanos and Vila (2009), the interest rates of all maturities are determined through the interaction between risk-averse arbitrageurs and investor clienteles with preferences for specific maturities. In this framework, the central bank purchases of long-term Treasuries can lower the long-term yields because they create a “scarcity effect” that arbitrageurs cannot eliminate. Moreover, the purchases can be effective, as they shorten the average maturity of government debt and therefore the duration risk held by arbitrageurs.

In this paper we investigate the effects of the ECB’s purchases of bank covered bonds and euro-zone sovereign debt. These assets are more risky than government bonds considered in Vayanos and Vila (2009) and the duration risk is not the only one that the central bank takes on its balance sheet. By purchasing the above-mentioned assets, the ECB also accepts the liquidity and default risk that private investors do not want to hold and replaces it with risk-free reserves. Private investors can ask for smaller liquidity compensation when buying covered or sovereign bonds if they know that they will be able to easily sell the assets to the ECB.⁷

⁷De Pooter, Martin, and Pruitt (2012) build a structural search-based asset pricing model that accounts for default risk in Europe and gives rationale for the ECB sovereign bond purchases.

The sovereign debt crisis in Europe increased default risk in the sovereign bond markets. Market participants started to price in not only high probability of sovereign default but also the probability that some member states would exit the euro zone. Such projections cut off these countries' access to market refinancing or made it extremely costly, leading to a "self-fulfilling" prophecy: default or exit of a country from the euro zone. By purchasing government bonds, and indirectly securing the sovereign debt, the ECB intended to prevent this "bad equilibrium" outcome.

There exists another transmission channel of central bank asset purchases which, instead of reducing risk premia, has an impact on the private sector's expectations of the future monetary policy ("signaling effect").⁸ In this paper, however, we focus on the ECB's impact on risk premia rather than on agents' expectations of future monetary policy, given that the ECB objective was to restore homogeneous credit conditions throughout the euro zone but not necessarily to ease credit conditions in aggregate (Coeuré 2012). Increased spreads on covered and sovereign bond markets in some member countries were a reflection of these divergent credit conditions in the euro zone.

2.2.1 Sovereign Bond Purchases

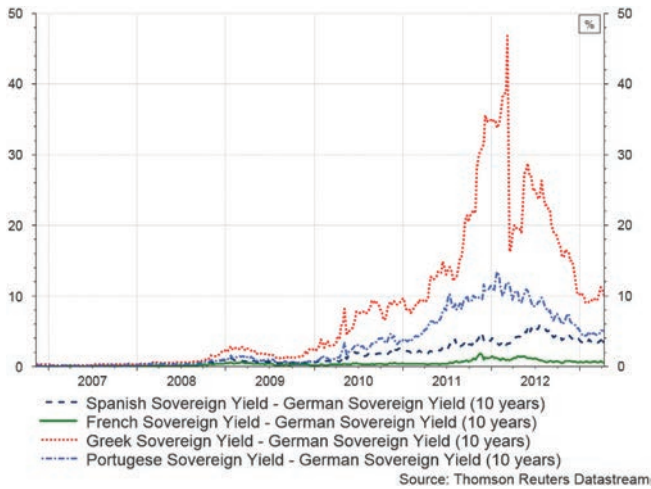
The Greek sovereign debt crisis in spring 2010 triggered a fire sale of some euro-zone government bonds. The ECB announced on Sunday, May 9, 2010 the Securities Market Programme (SMP) as a part of European Union efforts to stabilize the euro.⁹ The program was designed to purchase sovereign bonds and therefore to "ensure depth and liquidity in those market segments which are dysfunctional" (figure 3).¹⁰ The SMP was from the start a source of division within the ECB. Critics said that the ECB was overstepping

⁸Accumulation of risky assets on the central bank's balance sheet associated with important balance sheet expansion can be understood by financial markets as a signal that the monetary easing will continue longer than previously expected. Indeed, raising interest rates in these circumstances would expose the central bank to capital losses on the assets it holds.

⁹On the same day, the European Financial Stability Facility (EFSF) was established.

¹⁰The official ECB press release was issued on May 10, 2010: "ECB Decides on Measures to Address Severe Tensions in Financial Markets."

Figure 3. Sovereign Bond Spreads in the Euro-Zone Countries



its mandate by buying public debt in secondary markets and that the bond purchases would increase the inflationary pressures as well as undermine the ECB's credibility. However, the ECB insisted that the SMP was temporary and merely aimed at improving the transmission of the monetary policy. In order to distinguish the SMP from the U.S.-style quantitative easing and to ensure that the monetary policy stance was not affected, the ECB decided to sterilize these purchases via specific operations designed to reabsorb the injected liquidity.¹¹ Another notable difference between the SMP and the Federal Reserve sovereign bond purchases is that the ECB gave no details on the amount of bonds to be purchased, their origin, or how long it intended the program to last. The purchases stopped unofficially in January 2011, but the intensity of the euro-zone crisis and the risk of contagion to Italy and Spain made the ECB resume the program in August 2011. The ECB bought €219.5 billion of euro-zone government bonds within the SMP.

¹¹The sterilization of SMP seems mostly symbolic, as the FRFA procedure in all main refinancing operations leaves the control of monetary base in the hands of banks participating in these operations.

The euro-zone debt crisis continued in the beginning of 2012 as the critical financial standing of Spanish banks was revealed. The concerns about their solvency and *in fine* solvency of the Spanish government made the sovereign yields in the euro-zone periphery increase rapidly, as market participants were pricing in the possibility of some countries leaving the Monetary Union. As a response, ECB President Mario Draghi announced in July 2012 that the central bank would do “whatever it takes to save the euro.”¹² On September 6, 2012, the ECB announced the sovereign bond purchasing program—Outright Monetary Transactions (OMT)—and at the same time officially terminated the SMP. The objective of the new program, like the objective of the SMP, was to repair the monetary policy transmission mechanism and restore homogeneous credit conditions throughout the euro zone. More precisely, the purchases of euro-zone periphery sovereign debt were intended to reduce the risk premia related to fears of the reversibility of the euro. Despite the shared objective, the OMT was different from the SMP in several aspects. First, the maximum maturity of bonds purchased was set to three years, whereas the SMP concerned longer-term bonds. Second, there was a conditionality attached to participating in the OMT: the ECB would only purchase sovereign debt of a given country if its government complied with a full or precautionary macroeconomic adjustment program set by the European Financial Stability Facility (EFSF) or the European Stability Mechanism (ESM). Third, the ECB decided to forgo its seniority status with respect to private creditors. Finally, once the country met the access conditions, the ECB would intervene without limits, whereas the SMP was always presented as “temporary” and “limited,” which was hardly reassuring for investors.¹³ The ECB has not purchased any sovereign bonds within OMT since the announcement of the program.

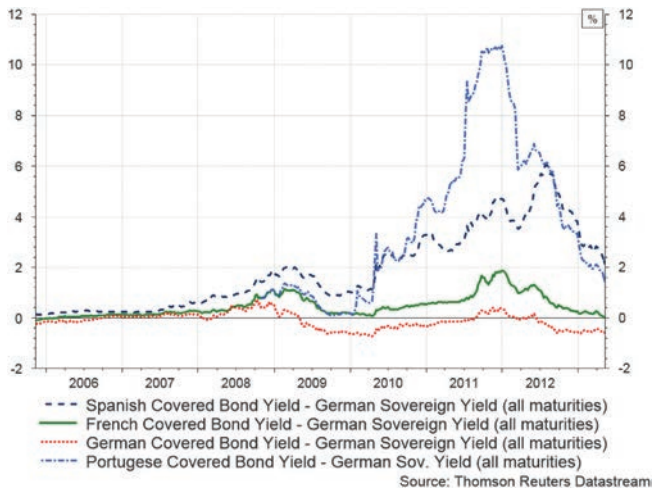
2.2.2 Covered Bond Purchases

Covered bonds are securities issued by credit institutions to assure their medium- and long-term refinancing. They are collateralized by

¹²See Draghi (2012).

¹³See the introductory statement to the ECB press conference, November 3, 2011, available at <http://www.ecb.int/press/pressconf/2011/html/is111103.en.html>.

Figure 4. Covered Bond Spreads in the Euro-Zone Countries



a dedicated pool of loans, typically mortgage loans and public-sector loans, and remain on the lender's balance sheet. They are seen as safer than other bank bonds, because they give investors a claim on the credit institution itself and on the cover pool of collateral as well. At the end of 2007 covered bonds were the most important privately issued bond segment in Europe's capital markets (ECB 2008). Despite their initial resilience to the financial turmoil that started in August 2007, this market dried up after Lehman Brothers collapsed in September 2008, as investors turned to government bonds and other less risky assets (figure 4). To prevent a credit crunch, the ECB announced on May 7, 2009 that it would purchase €60 billion of euro-denominated covered bonds issued in the euro zone. This decision was surprising for the markets which were expecting the rate cut and the lengthening of the lending program but not the purchases of private debt, which were perceived as a change in strategy.¹⁴ The objectives of the Covered Bond Purchase Programme (CBPP1) were the following: promoting the ongoing

¹⁴From "Trichet Drags ECB Into New Era Over Weber's Bond Objections," Bloomberg, May 7, 2009.

decline in money-market term rates; easing funding conditions for credit institutions and enterprises; encouraging credit institutions to maintain and expand their lending to clients; and improving market liquidity in important segments of the private debt securities market.

At the end of June 2010, the ECB stopped the covered bond purchases, but as the sovereign crisis deepened in autumn 2011, it proceeded to further measures supporting the covered bond markets. On October 6, 2011 it announced the second Covered Bond Purchase Programme (CBPP2) of €40 billion in favor of euro-denominated covered bonds in both primary and secondary markets.

3. Methodology

We apply event-based regressions to measure the impact of the ECB's unconventional monetary policies on bank and government borrowing costs. The borrowing conditions for banks are measured by money-market spreads (short-term funding) and covered bond spreads (longer-term funding). Government borrowing costs are approximated by sovereign bond spreads. Event-based regressions allow testing the impact of an economic event on financial market data. In modern financial markets, such as those in the euro zone, the effect of an event should be reflected in asset prices over a short period of time.

One should note, however, that while the event-based regressions allow measuring the impact of the ECB measures on announcement days, they do not indicate whether these effects were persistent. Wright (2012) employs the VAR methodology and shows that the effects of the quantitative easing (QE) implemented by the Federal Reserve wear off over the next few months. There are several possible reasons for this yield behavior. First, QE might have triggered higher growth and inflation expectations that would finally increase rates. Second, and more relevant in the euro zone, the initial effect of the announcement could be offset if the markets subsequently learn that the measure may not be fully implemented. For instance, it is plausible that the support of the German chancellor for the OMT increased the credibility of the announcement, while the public objections of Axel Weber to the SMP contributed to the revision

of the markets' initial assessment of the measure.¹⁵ Other factors, such as subsequent news about deficit, fiscal consolidation, euro-zone bailout facilities, or banking union agreements, can contribute to the reversal or reinforcement of the ECB measures.¹⁶ Even though the evaluation of the long-term effectiveness of the ECB measures cannot be accomplished with the event-based methodology, we believe that their short-term impact on financial markets is still important. During the crisis, even temporary relief to bank and government funding markets was crucial for the euro-zone financial stability.

We use daily data from July 2, 2007 until September 27, 2012, with the exception of Italian and Portuguese covered bond series available, respectively, from January 2, 2009 and October 31, 2008. We rely on dummy variables to discriminate between days when announcements were made or not. Based on the ECB press releases, we create a database of unconventional monetary policy news. The announcements are classified into seven categories described in the previous section:

- Exceptional liquidity measures:
 - Fixed-rate full-allotment procedure (FRFA)
 - Three-year LTROs: announcement (three-year LTRO)
 - Three-year LTROs: operation dates (three-year LTRO op.)
 - 0 percent deposit rate at the ECB (0% deposit)
- Asset purchases:
 - Covered bond purchase programs (CBPP1 and CBPP2)
 - Longer-term sovereign bond purchase program (SMP)
 - Short-term sovereign bond purchase program (OMT)

¹⁵From "Doubts on ECB Bond Buying," *Financial Times*, June 8, 2010: "The effectiveness of the bond purchasing programme has certainly been undermined by comments from Axel Weber, Germany's Bundesbank president, who has called for a tight limit on the amount of bonds to be bought."

¹⁶From "Burden Switches to EU Politicians," *Financial Times*, September 7, 2012 (after the OMT announcement): "The stunning two-day rally in euro-zone debt markets after Thursday's European Central Bank decision to overhaul and restart its sovereign bond buying plan is not the first time traders have cheered what appeared to be decisive action by European Union policymakers. In the past, however, that optimism has often been followed by sharp sell-offs triggered by rating downgrades, bungled communication or leaders failing to follow through on the policies they had promised."

In the empirical section we estimate the following regressions:

$$\begin{aligned} \Delta S_t^{M,C,S} = & \alpha + \sum_{i=1}^7 \beta_i NC_{i,t} + \varphi_1 F_t + \varphi_2 C_t \\ & + \sum_{n=1}^4 \gamma_n \Delta S_{t-n}^{M,C,S} + \sum_{l=1}^4 \theta_l D_t + \varepsilon_t, \end{aligned} \quad (1)$$

where the dependent variable is the money-market spread, ΔS_t^M (section 4.1); covered bond spread, ΔS_t^C (section 4.2); and sovereign bond spread, ΔS_t^S (section 4.3). $NC_{i,t}$ are dummies for unconventional monetary policy announcements; F_t is a dummy for the European rescue funding program announcements (EFSF and ESM); C_t is a dummy for the sovereign debt crisis; $\Delta S_{t-n}^{M,C,S}$ are lagged values of dependent variable included to correct for the autocorrelations of the residuals ($n = 3$ for all money-market spreads except the Euribor-repo spread, for which $n = 4$; $n = 1$ for covered bond and sovereign spreads); D_t are dummies for the working day of the week (Monday, Tuesday ...); and ε_t is a stochastic error term.

The analysis of monetary policy announcements via event-based regressions presents several potential difficulties. First, the announcements studied must be unanticipated by the market participants (MacKinlay 1997). Otherwise, the impact of the event is incorporated before it is announced and there is no change in yields and prices on the announcement day. In other terms, only surprising events can be appropriately evaluated within this methodology. In practice, many of the unconventional ECB measures were anticipated by the markets. This is principally the case of supplementary liquidity announcements, such as lengthening the maturity of the refinancing operations up to one year, which are quite close to conventional liquidity provisions and do not imply much risk on the ECB balance sheet. For that reason we focus our analysis only on the major unconventional policy announcements that were surprising and important news to the markets.¹⁷ We define unexpected announcements based on the qualitative content of the news flow

¹⁷As a robustness check, we tested the impact of smaller supplementary liquidity announcements but did not find any significant effect. We explain this issue in the next section.

in the financial press that followed and preceded each measure. It is worth noting that the majority of the measures considered were announced on unscheduled ECB Governing Council meetings, rendering the anticipation of the measure unlikely. Moreover, the unconventional toolbox being quite large, the exact choice of the measure was more difficult to anticipate. Yet, some of the measures were partly anticipated, which is the case of the three-year LTROs and OMT. We decided to include them in the analysis given their novelty. The strong reaction of the interbank market in the case of the three-year LTROs and of the covered and sovereign bond markets in the case of OMT confirms that these events still incorporated some elements of surprise.¹⁸ In the case of leakage before the announcement, our findings constitute a lower bound for the effectiveness of the ECB measures.

The second issue linked to event-based regressions is the simultaneous occurrence of other events on the day of monetary announcement that might affect the variables of interest and therefore bias the results. It seems particularly important during the crisis, when there were sometimes several policy measures announced on the same day. When these events coincide with monetary policy announcements, it is necessary to include them in regression in order to separate their effects. We use the Factiva press database to check if there were other major events that might have influenced our variables of interest, i.e., interest rate spreads.¹⁹ The most striking example of simultaneous announcements is the weekend of May 8–9, 2010 when several monetary measures were decided and, in particular, the SMP was created. In parallel, the euro-zone politicians founded the European Financial Stability Fund (EFSF). Even though both the SMP and the EFSF were intended to purchase sovereign debt, it is useful to separate the effects of the two measures, as they are conducted by different institutions. To assure a correct specification of our event-based regression model, we include announcements concerning the EFSF and the ESM developments as well as the

¹⁸Even though some market participants evoke the possibility of an unconventional measure announcement, the remaining uncertainty does not allow a full pricing of the event until the actual announcement takes place.

¹⁹Factiva is an information and research tool owned by Dow Jones & Company. It offers online articles from both licensed and free sources (*Wall Street Journal*, Reuters, and *Financial Times*, among others).

dummy for the sovereign debt crisis.²⁰ The crisis dummy is equal to 1 during the periods when the concerns about solvency of the periphery euro-zone countries were the highest.²¹ Additionally, in case the ECB announcements and the spreads react to unobserved news, we control for this factor by including the VIX volatility index.

Finally, there might be a simultaneous causal relationship between the ECB's unconventional policy announcements and the financial market spreads. However, we believe that this is mainly a problem when the actual operations are being evaluated and is negligible when the announcement effect is considered. For instance, the ECB's asset purchases were most likely on the days when the sovereign spreads increased. However, it is not plausible that the ECB announced their important unconventional measures based on the previous day's change in spreads. These measures are often politically contested and involve a long decision process.

4. Results

4.1 Money Market

Since August 2007 the spreads between unsecured and secured rates on the interbank market increased to previously unseen levels (figure 2). The ECB was determined to support money-market activity, as interbank lending is a key element of the successful monetary transmission. The three-year LTROs, FRFA procedure, covered bond purchase programs, and setting the deposit rate to zero favored interbank lending by reducing funding constraints and increasing the cost of liquidity holding. Sovereign bond purchase programs could also have a positive effect on money markets, given that banks were holding large quantities of euro-area sovereign debt.

To test the impact of unconventional policy measures on the money-market spreads, we estimate equation (1), where the dependent variable ΔS_t^M is a two-day change in three-month

²⁰We checked several other categories of announcements that might affect the spreads, such as European Summits or conventional monetary policy surprises, but they did not occur on the same days as unconventional monetary policy announcements.

²¹A detailed description of all dummies used in the regressions can be found in table 5 in the appendix.

money-market spreads. We use four alternative measures of money-market distress reflecting the difference between unsecured and secured (or risk-free) three-month interest rates: (i) the Euribor-OIS spread,²² (ii) the Euribor-repo spread,²³ (iii) the Euribor-German Treasury bill spread, and (iv) the certificate of deposit (CD)-OIS spread.²⁴ Among these measures, the Euribor-OIS is the most commonly cited barometer of interbank market tensions. The recent revelations about LIBOR and Euribor manipulation by one of the contributing banks cast some doubts on its credibility. However, there are two particular features of the Euribor rate that make it less sensitive to manipulation than LIBOR. First, forty-three banks contribute to Euribor, as opposed to fifteen in the Euro Libor panel, which reduces the weight of the eventual misreporting contributor. Second, Euribor is an average lending rate, while LIBOR is an average borrowing rate. During a crisis, the contributing banks are more inclined to diminish the latter, as high borrowing rates send a negative signal about their financial standing.

There is also a timing issue related to the Euribor-OIS spread. The Euribor rate is published at 11:00 a.m. Brussels time (10:00 GMT), while the last update of the OIS rate in Datastream is done at 19:15 GMT. Therefore, many announcements on a given day are not taken into account by the Euribor rate. In order to ensure that the markets had the possibility to react to all announcements, we consider a two-day event window: the change in spread one day after the announcement with respect to the day before ($\Delta S_t^M = S_{t+1}^M - S_{t-1}^M$).

Table 1 reports the estimation results. Among all unconventional ECB policies, reducing the ECB deposit rate to 0 percent and announcing three-year LTROs diminished money-market spreads

²²The Euro Interbank Offered Rate (Euribor) is an average interbank borrowing rate published daily at 11:00 a.m. (Brussels time) by the European Banking Federation (EBF). The overnight index swap (OIS) rate represents market expectations of the monetary policy rate over the future months. There is no exchange of principal and only the net difference in interest rates is paid at maturity, so there is very little default risk in the OIS market.

²³The repo rate is the rate at which, at 11:00 a.m. Brussels time, one bank offers (in the euro zone and worldwide) funds in euros to another bank if, in exchange, the former receives from the latter the best collateral within the most actively traded European repo market.

²⁴A certificate of deposit is a debt instrument issued by banks and other financial institutions.

Table 1. Impact of the ECB's Unconventional Monetary Policies on the Money Market

This table presents regression results for the euro-zone money market: $\Delta S_t^M = \alpha + \sum_{i=1}^7 \beta_i N C_{i,t} + \varphi_1 F_t + \varphi_2 C_t + \sum_{n=1}^4 \gamma_n \Delta S_{t-n}^M + \sum_{l=1}^4 \theta_l D_t + \varepsilon_t$. Dependent variables are two-day changes in money-market spreads: Euribor-OIS (three-month Euribor rate minus three-month overnight index swap rate), Euribor-repo (three-month Euribor rate minus three-month repo rate), Euribor-Bubill (three-month Euribor rate minus three-month German Treasury bill), and CD-OIS (three-month certificate of deposit rate minus three-month overnight index swap rate). Independent variables are dummy variables: EFSF/ESM = 1 when important developments about the EFSF and the ESM are announced (May 10, 2010; March 14, 2011; March 26, 2012; June 29, 2012); SMP = 1 when longer-term sovereign bond purchases are announced (May 10, 2010); OMT = 1 when short-term sovereign bond purchases are announced (September 6, 2012); CBPP1 and CBPP2 = 1 on May 7, 2009 and October 6, 2011; three-year LTRO announcement = 1 on December 8, 2011; three-year LTRO operations = 1 on December 21, 2011 and February 29, 2012; FRFA = 1 on October 9, 2008; October 13, 2008; October 15, 2008; and May 10, 2010; 0% deposit rate at the ECB = 1 on July 5, 2012. Long-term coefficients are reported: $\text{coeff}/(1 - \sum_{n=1}^4 \gamma_n)$. Constant, day dummies, and lags of dependent variables are included but not reported.

	Euribor-OIS	Euribor-Repo	Euribor-Bubill	CD-OIS
Sovereign Crisis Dummy	0.00 [0.55]	0.00 [0.30]	0.00 [0.61]	0.00 [0.41]
EFSF/ESM	0.02 [0.28]	0.02 [0.15]	0.02 [0.64]	-0.01 [0.89]
SMP	0.27 [0.31]	0.03 [0.89]	-0.27 [0.19]	0.07 [0.62]
OMT	-0.02*** [0.00]	-0.02*** [0.00]	-0.01* [0.08]	
CBPP1 and CBPP2	-0.07*** [0.00]	-0.12 [0.42]	-0.06*** [0.00]	-0.04 [0.14]
Three-Year LTRO Announcement	-0.22*** [0.00]	-0.14*** [0.00]	-0.06*** [0.00]	
Three-Year LTRO Operations	-0.06*** [0.03]	-0.05*** [0.03]	-0.03*** [0.00]	-0.12*** [0.00]
FRFA	-0.32 [0.24]	-0.03 [0.89]	0.22 [0.26]	-0.04 [0.79]
0% Deposit Rate at the ECB	-0.24*** [0.00]	-0.10*** [0.00]	-0.14*** [0.00]	
Observations	1,365	1,364	1,278	1,187
R-squared	0.48	0.61	0.33	0.21
Notes: Robust p-values are in brackets. *** = $p < 0.01$, ** = $p < 0.05$, and * = $p < 0.1$.				

the most significantly. The Euribor-OIS spread fell by 24 basis points after the ECB decided to lower its deposit rate to 0 percent, while the Euribor-repo and Euribor-German Treasury bill spreads fell by, respectively, 10 and 14 basis points.²⁵ The announcement of three-year LTROs reduced the Euribor-OIS spread by 22 basis points, the Euribor-repo spread by 14 basis points, and the Euribor-German Treasury bill spread by 6 basis points. Interestingly, the spreads fell also on the days of the three-year LTROs, even though market participants knew the operation dates in advance. The effect of operations was smaller and less significant than the announcement effect for the Euribor spreads (3 to 6 basis points) but reached 12 basis points for the CD-OIS spread. The FRFA procedure, on the other hand, did not reduce the spreads. These results confirm that three-year operations were indeed exceptional measures and incomparable to other liquidity facilities which were found ineffective in reducing money-market tensions. In particular, Angelini, Nobili, and Picillo (2011) find that liquidity risk was not the most important determinant of the spreads and therefore liquidity measures, such as LTROs up to one year, were not able to affect them.²⁶ Our results show that in the presence of high credit risk, stronger ECB measures such as three-year LTROs, cutting the deposit rate to 0 percent, or bank debt asset purchases can lower the spreads.

We find that bank covered bond purchases (CBPP1 and CBPP2) also diminished the spreads, but the magnitude of the results is

²⁵The coefficient is not reported for the CD-OIS spread, as there was no quotation for the three-month certificate of deposit on the day of the announcements. We use Reuters time series for the three-month certificate of deposit and German Treasury bill, while the Euribor, repo, and OIS rates come from Datastream.

²⁶This result might be also explained by the “crowding-out” effect of the ECB liquidity interventions as shown in Brunetti, di Filippo, and Harris (2011). The unlimited liquidity available at the central bank impairs the important functions of interbank transactions such as information aggregation, price discovery, and peer monitoring. However, shorter-term liquidity measures became quite common during the crisis and they might have been anticipated by the markets, which could also explain the lack of the effect on the announcement days. The difficulty of identifying the surprise component of smaller liquidity measures prevented us from including them in our analysis. Nevertheless, we conducted tests with announcements of lengthening the maturity of LTROs up to one year and—like Angelini, Nobili, and Picillo (2011) and Brunetti, di Filippo, and Harris (2011)—found no significant impact of these announcements, while other results remained valid.

smaller. The significant effects range from 6 to 7 basis points for the Euribor-OIS and Euribor-German Treasury bill spreads. On the other hand, sovereign bond purchases had either a very small impact (OMT) or no impact at all (SMP).

We conducted additional tests including the variables that drive money-market spreads in normal times, such as the risk-aversion proxy, VIX.²⁷ The results do not change significantly, with the exception of the impact of the OMT, which becomes positive and insignificant once the VIX is added.²⁸ In general, adding financial variables that determine spreads in normal times, such as the VIX, weakens slightly the impact of unconventional monetary policies, as these policies also have an impact on the VIX. Given the robustness of the results under both specifications, we decided to use as a benchmark the regression without the VIX.

4.2 Covered Bond Market

While the money market provides banks with short-term funding, the covered bond market is one of their longer-term funding sources. Despite its relative soundness in the beginning of the crisis, this market also dried up after the Lehman Brothers collapse, leading to unusually high risk premia (figure 4). Covered bond purchase programs were the main unconventional policy designed to reduce the covered bond spreads. However, other ECB unconventional measures could also reduce the cost of refinancing on this market. First, three-year LTROs and the FRFA procedure were intended to diminish bank funding risk and therefore encourage investors to ask for smaller risk compensation on bank debt instruments, such as covered bonds. Second, sovereign bond purchases—by improving the balance sheets of banks holding sovereign assets—could increase their creditworthiness and diminish covered bond spreads.

The ultimate objective of the ECB's asset purchases was to improve the monetary policy transmission and to restore homogeneous credit conditions across the euro-area member countries. Therefore, the effectiveness of the ECB measures is reflected not only by the overall euro-zone spread decrease but also by the particular

²⁷VIX is a measure of the implied volatility of S&P 500 index options.

²⁸The results with the VIX variable are available upon request.

impact in the periphery euro-zone countries. This is why we test the impact of the ECB policies not only on the synthetic euro-zone spread but also on the individual spreads in the core and periphery euro-zone countries.

We measure the impact of the ECB's unconventional monetary policies on covered bond spreads by estimating equation (1), where the dependent variable ΔS_t^C is a one-day change in covered bond spread in the euro zone and in its member countries: France, Germany, Ireland, Italy, Portugal, and Spain.²⁹ The spreads are calculated as the difference between covered bond yields in each country with respect to the corresponding German sovereign bond yield. All bond rates are synthetic benchmark yields provided by iBoxx and available from Datastream. They cover all maturities exceeding one year and are comparable among countries. The composed-maturity bond indexes seem appropriate, as the ECB bought covered bonds of different maturities.³⁰

Table 2 presents the estimation results for the euro zone, France, Germany, Ireland, Italy, Portugal, and Spain. At the euro-zone level, long-term sovereign bond purchases (SMP) had the strongest impact and diminished the covered bond spread by 19 basis points (bp). They were followed by short-term sovereign bond purchases, OMT (−6 bp), covered bond purchases (−4 bp), and three-year LTROs (−4 bp). The positive news concerning the European rescue funding programs (EFSF/ESM) also diminished spreads (−4 bp), while the sovereign crisis dummy increased it (+1 bp).

It is interesting to note that sovereign bond purchases have a bigger impact on covered bond spreads than covered bond purchases themselves. It seems that the sovereign risk in the euro area and the fact that the banks hold sovereign bonds play an important role in the ECB's unconventional policy transmission. Breaking up the results by country confirms this intuition.

Table 2 shows that the SMP had a much stronger effect in troubled periphery countries (Portugal, −163 bp; Ireland, −46 bp; Spain, −34 bp; Italy −31 bp) than in core euro-zone countries (Germany, −10 bp; France, −7 bp). The second sovereign bond purchasing

²⁹Datastream does not provide the iBoxx covered bond rates for Greece.

³⁰CBPP1: 3–10 years, with strong focus on maturities up to 7 years; CBPP2: Up to 10.5 years residual maturity, according to the ECB website.

Table 2. Impact of the ECB’s Unconventional Monetary Policies on the Covered Bond Market

This table presents regression results for the euro-zone covered bond market: $\Delta SC_t = \alpha + \sum_{i=1}^7 \beta_i NC_{i,t} + \varphi_1 Ft + \varphi_2 Ct + \gamma_n \Delta SC_{t-1} + \sum_{l=1}^4 \theta_l D_l + \varepsilon_t$. Dependent variables are one-day changes in covered bond spreads in the euro zone and its member countries: France, Germany, Ireland, Italy, Portugal, and Spain. The spreads are calculated as the difference between all-maturities covered bond yields for each country with respect to the corresponding German sovereign bond yield. Independent variables are dummy variables: EFSF/ESM = 1 when important developments about the EFSF and ESM are announced (May 10, 2010; March 14, 2011; March 26, 2012; and June 29, 2012); SMP = 1 when longer-term sovereign bond purchases are announced (May 10, 2010; 0% deposit rate at the ECB = 1 when short-term sovereign bond purchases are announced (September 6, 2012); CBPP1 and CBPP2 = 1 on May 7, 2009 and October 6, 2011; three-year LTRO announcement = 1 on December 8, 2011; three-year LTRO operations = 1 on December 21, 2011 and February 29, 2012; FRFA = 1 on October 9, 2008; October 13, 2008; October 15, 2008; and May 10, 2010; 0% deposit rate at the ECB = 1 on July 5, 2012. Long-term coefficients are reported: $\text{coeff}/(1 - \gamma)$. Constant, day dummies, and lags of dependent variables are included but not reported.

	Euro Zone	France	Germany	Ireland	Italy	Portugal	Spain
Sov. Crisis	0.01*** [0.00]	0.01*** [0.00]	0.00** [0.04]	0.01*** [0.00]	0.02*** [0.00]	0.04*** [0.00]	0.02*** [0.00]
EFSF/ESM	-0.04*** [0.00]	-0.02*** [0.00]	-0.03*** [0.00]	-0.06*** [0.00]	-0.06*** [0.00]	-0.12* [0.07]	-0.08*** [0.00]
SMP	-0.19*** [0.00]	-0.07*** [0.01]	-0.10*** [0.01]	-0.46*** [0.00]	-0.31*** [0.00]	-1.63*** [0.00]	-0.34*** [0.00]
OMT	-0.06*** [0.00]	-0.04*** [0.00]	-0.05*** [0.00]	-0.10*** [0.00]	-0.11*** [0.00]	-0.49*** [0.00]	-0.12*** [0.00]
CBPP1 and CBPP2	-0.04*** [0.00]	-0.04*** [0.00]	-0.05*** [0.02]	0.04 [0.57]	-0.10*** [0.00]	-0.09 [0.26]	-0.06*** [0.00]
Three-Year LTRO An.	-0.04*** [0.00]	-0.05*** [0.00]	-0.02*** [0.00]	-0.05*** [0.00]	-0.03*** [0.00]	0.03 [0.78]	-0.03*** [0.00]
Three-Year LTRO Op.	-0.01 [0.34]	-0.01 [0.20]	-0.01 [0.77]	-0.01 [0.68]	-0.04 [0.17]	0.00 [0.97]	-0.01 [0.31]
FRFA	-0.04 [0.14]	-0.03 [0.22]	-0.03 [0.42]	-0.04 [0.47]	-0.04 [0.01**]	-0.06** [0.03]	-0.06** [0.03]
0% Deposit	-0.00 [0.10]	-0.02*** [0.00]	-0.01*** [0.00]	-0.01 [0.28]	0.01** [0.03]	-0.10 [0.38]	0.07*** [0.00]
Observations	1,368	1,368	1,368	1,368	973	1,018	1,368
R-squared	0.13	0.08	0.06	0.13	0.17	0.27	0.19

Notes: Robust p-values are in brackets. *** = $p < 0.01$, ** = $p < 0.05$, and * = $p < 0.1$.

program (OMT) also had the most important effect in the euro-zone periphery. These results suggest that sovereign bond purchases have important effects on longer-term bank debt instruments when sovereign risk is high. By diminishing sovereign risk in periphery countries, these purchases improve the credit standing of the financial institutions and increase the price of their covered bonds. Indeed, banks in these troubled countries hold an important amount of sovereign debt, and their own creditworthiness depends largely on the prices of the sovereign assets. These results indicate therefore the connection between bank and sovereign default risk: by diminishing sovereign default risk, the ECB managed to reduce the risk compensation for the bank debt instruments. The reduction of covered bond spreads after the announcements of European sovereign rescue facilities (EFSF/ESM), also designed to purchase sovereign debt, confirms that investors in bank covered bonds were sensible to measures reducing the probability of sovereign default. The presence of the sovereign-bank feedback loop seems to matter for the ECB's asset purchases: *sovereign* bond purchases diminish bank *covered* bond spreads. In the next section we will show that the opposite is also true: purchases of bank *covered* bonds diminish *sovereign* bond spreads.

The measures directly aimed at relieving bank funding constraints were also effective but to a smaller extent. As expected, covered bond purchase programs (CBPP1 and CBPP2) diminished covered bond spreads in all countries studied, with the exception of Ireland and Portugal: France (−4 bp), Germany (−5 bp), Italy (−10 bp), and Spain (−6 bp).³¹ These results are not surprising, given that the biggest amounts of CBPP1 were allocated to the central banks of France, Germany, Italy, the Netherlands, and Spain (European Covered Bond Council 2010). Furthermore, Finland and Italy were the main beneficiaries when the ratio of purchased amounts to the size of the outstanding covered bonds eligible under the CBPP1 is taken into account. In contrast to the sovereign bond purchases, however, covered bond purchases had quantitatively similar effects

³¹We tested CBPP1 and CBPP2 separately and they both have a similar impact on covered bond spreads, with the exception of Ireland, for which the first program increased the spreads and the second diminished it. Beirne et al. (2011) also find that CBPP1 did not reduce the Irish covered bond spreads.

in periphery and core euro-zone countries. They were not addressed to reduce the sovereign risk and worked through the more traditional “scarcity channel.”

Three-year LTROs had an impact comparable to that of covered bond purchases. They reduced spreads in France (−5 bp), Germany (−2 bp), Ireland (−5 bp), Italy (−3 bp), and Spain (−3 bp). Three-year LTROs reduced longer-term bank funding constraints and therefore diminished bank liquidity and credit risk, pulling the yield on bank debt down. The FRFA procedure also contributed to spread reduction, particularly in Spain (−6 bp).

Overall, our results show that measures designed to reduce the probability of sovereign default, such as sovereign bond purchases, diminish the most bank covered bond spreads. Measures aiming at reducing directly the bank funding cost were also effective, albeit to a smaller extent, in lowering the spreads and had a homogeneous effect across member states.

4.3 Sovereign Bond Market

The sovereign yields in the periphery euro-zone countries increased dramatically with respect to corresponding German yields since the sovereign debt crisis started (figure 3). To ensure the depth and liquidity of these markets, the ECB announced two government bond purchasing programs (SMP and OMT). Moreover, other ECB policies, such as those addressed to banks (three-year LTROs, the FRFA procedure, or covered bond purchase programs), could in principle have a positive impact on sovereign spreads. Indeed, improving bank creditworthiness diminishes the probability of sovereign default by reducing the potential necessity of bank bailout by a government. Furthermore, granted with abundant liquidity, banks themselves could purchase sovereign bonds and therefore increase their prices.

We measure the effects of the ECB’s unconventional policies on sovereign spreads in periphery euro-zone countries and in France. Just as in the case of covered bond spreads, the effectiveness of the ECB policies should be reflected by more homogeneous sovereign spreads across the euro zone after the ECB intervention. In other terms, the ECB measures should have a bigger impact on the troubled periphery countries. To this end, we estimate equation (1), where the dependent variable ΔS_t^S is a one-day change in the

ten-year sovereign bond spread. The spread is calculated as the difference between the ten-year sovereign yields in the euro zone or its member countries (France, Greece, Ireland, Italy, Portugal, and Spain) and the ten-year German sovereign yield. The ten-year euro-zone benchmark yield is available from Datastream and represents a weighted average of bond yields from each euro-zone member.³²

Table 3 presents the results for the euro zone, France, Greece, Ireland, Italy, Portugal, and Spain. The most striking result in the euro zone is the impact of the ECB longer-term sovereign bond purchase program (SMP), which reduced the spreads by 16 basis points. The effect is particularly strong for the countries where sovereign risk attained the highest levels: Greece (−485 bp), Ireland (−121 bp), and Portugal (−202 bp). Italy and Spain acknowledge the reduction of, respectively, 31 and 43 basis points, while the French spread does not react to the SMP announcement. The SMP program was announced in the midst of the Greek debt crisis without any precision about the amounts or the regularity of the purchases. Every Monday, the ECB released information about the amount of bonds purchased the previous week but did not mention their country of origin. According to some analysts, in the first phase of the SMP the ECB purchased Greek, Irish, and Portuguese bonds, and this is reflected in our regression results.³³

As the euro sovereign debt crisis was about to spread to Italy and Spain, the ECB confirmed its willingness to purchase actively the euro-zone sovereign bonds. This announcement was made on Sunday, August 7, 2011 and was preceded by a positive appreciation of the Italian and Spanish austerity program execution. It was unambiguously understood by market participants as a promise to buy Italian and Spanish government bonds. We take this announcement into account (SMP2) and report the results in table 4. The overall effect on the benchmark euro-zone spread is significant (−26 bp) and, as expected, the biggest impact is observed for Spain (−104 bp) and Italy (−82 bp).

³²The weightings used are the 1996 real GDP as published by Eurostat.

³³From “ECB Keeps Bond-Buying Programme Dormant,” Reuters, August 1, 2011: “Bond market traders and analysts say buying has been limited to Greek, Irish and Portuguese bonds and estimate that it holds around 45 billion euros of Greek debt.”

Table 3. Impact of the ECB’s Unconventional Monetary Policies on the Sovereign Bond Market

This table presents regression results for the euro-zone sovereign bond market: $\Delta S_t^S = \alpha + \sum_{i=1}^9 \beta_i N C_{i,t} + \delta Q_t + \varphi_1 F_t + \varphi_2 \hat{C}_t + \sum_{n=1}^2 \gamma_n \Delta S_{t-n}^S + \sum_{l=1}^4 \theta_l D_l + \varepsilon_t$. Dependent variables are one-day changes in sovereign bond spreads in the euro-zone member countries. The spreads are calculated as the difference between ten-year sovereign bond yields for each country with respect to the corresponding German sovereign yield. Independent variables are dummy variables: EFSF/ESM = 1 when important developments about the EFSF and the ESM are announced (May 10, 2010; March 14, 2011; March 26, 2012; June 29, 2012); SMP = 1 when longer-term sovereign bond purchases are announced (May 10, 2010); OMT = 1 when short-term sovereign bond purchases are announced (September 6, 2012); CBPP1 and CBPP2 = 1 on May 7, 2009 and June 10, 2011; three-year LTRO announcement = 1 on December 8, 2011; three-year LTRO operations = 1 on December 21, 2011 and February 29, 2012; FRFA = 1 on October 9, 2008; October 13, 2008; October 15, 2008; and May 10, 2010; 0% deposit rate at the ECB = 1 on July 5, 2012. Long-term coefficients are reported: $\text{coeff}/(1 - \sum_{n=1}^2 \gamma_n)$. Constant, day dummies, and lags of dependent variables are included but not reported.

	Euro Zone	France	Greece	Ireland	Italy	Portugal	Spain
Sov. Crisis	0.01** [0.05]	0.00 [0.17]	0.13*** [0.00]	0.02*** [0.01]	0.01** [0.05]	0.02 [0.23]	0.02** [0.01]
EFSF/ESM	-0.13*** [0.00]	-0.07*** [0.00]	-0.25* [0.08]	-0.52*** [0.00]	-0.28** [0.01]	-0.46*** [0.00]	-0.43*** [0.00]
SMP	-0.16*** [0.00]	-0.03 [0.48]	-4.85*** [0.00]	-1.21*** [0.00]	-0.31*** [0.01]	-2.02*** [0.00]	-0.43*** [0.00]
OMT	-0.14*** [0.00]	-0.07* [0.06]	0.05 [0.30]	-0.24* [0.09]	-0.31*** [0.00]	-0.54*** [0.01]	-0.59*** [0.00]
CBPP1 and CBPP2	-0.05*** [0.00]	-0.01 [0.76]	-0.18*** [0.01]	-0.09 [0.39]	-0.16*** [0.00]	-0.08 [0.58]	-0.13* [0.10]
Three-Year LTRO An.	0.18*** [0.00]	0.18*** [0.00]	1.21*** [0.00]	0.05 [0.73]	0.48*** [0.00]	0.09 [0.68]	0.33*** [0.00]
Three-Year LTRO Op.	-0.00 [0.95]	-0.03 [0.24]	-0.04 [0.83]	0.02 [0.82]	-0.01 [0.93]	0.38** [0.01]	0.05 [0.52]
FRFA	-0.02*** [0.01]	-0.02 [0.43]	-0.00 [0.94]	-0.01 [0.95]	-0.04** [0.03]	-0.08 [0.53]	-0.04 [0.57]
0% Deposit	0.11*** [0.00]	0.02 [0.64]	-0.04* [0.06]	0.11 [0.45]	0.29*** [0.00]	0.33 [0.11]	0.52*** [0.00]
Observations	1,368	1,368	1,368	1,368	1,368	1,368	1,368
R-squared	0.13	0.06	0.04	0.19	0.10	0.19	0.18

Notes: Robust p-values are in brackets. *** = $p < 0.01$, ** = $p < 0.05$, and * = $p < 0.1$.

The second sovereign bond purchase program, OMT, had a similar impact on the benchmark euro-zone sovereign spread (-14 bp). The program was announced in a context of sovereign debt crisis in Spain, and the response of the Spanish spread to the OMT announcement was the strongest (-59 bp). The impact on the Italian and Portuguese spreads was also significant at 5 percent (respectively, -31 bp and -54 bp), while it was only at 10 percent for the Irish spreads and was not significant for the Greek spread. The French spread reacted only to a small extent (-7 bp, at the 10 percent significance level).

The OMT was announced on September 6, 2012, but some kind of ECB intervention in sovereign markets was expected since Mario Draghi's speech on July 26, 2012, in which he promised to do "whatever it takes to save the euro." We include this announcement in our regression and show the results in table 4. The French spread was reduced less (-12 bp) than the Spanish and Italian spreads after this announcement (respectively, -56 bp and -48 bp). The response of the Greek spread was significant and high (-56 bp), contrary to the actual OMT announcement later on. The July 26 announcement triggered expectations of long-term sovereign spreads targeting, but this feature was dismissed in the final version of the program, which might have disappointed potential investors on the Greek market.

The strong impact in the periphery euro zone suggests that central bank intervention in the sovereign market is particularly effective when sovereign risk is important. The fall of the sovereign bond spreads following the European rescue program announcements (EFSF/ESM, -13 bp) confirms that measures aimed at reducing sovereign default risk were effective in diminishing government borrowing costs in the euro zone.

Covered bond purchase programs (CBPP1 and CBPP2) were another measure that reduced the sovereign spreads (-5 bp). This result seems to reveal the second part of the sovereign-bank feedback loop we evoked in the previous section. Reducing bank-funding risk could indeed diminish sovereign risk, as it reduces the probability of future government-led bailout. Table 3 shows that the effect of CBPP1 and CBPP2 is particularly important in periphery euro-zone countries (Greece, -18 bp; Italy, -16 bp; Spain, -13 bp).

The puzzling result, however, is the reaction of the sovereign spreads following the important three-year LTROs announcement.

Contrary to expectations, sovereign spreads rise following this announcement, especially in the Southern European countries (table 3). The reaction of sovereign spreads to the three-year LTROs announcement is opposite to interbank market and covered bond market reactions, which were in line with expectations. This result indicates that three-year LTROs improved significantly market borrowing costs for the euro-zone banks but not governments. Given that three-year loans were granted to banks, this response may seem natural. However, the three-year LTROs announcement significantly *increased* government borrowing costs. This reaction suggests that there was another piece of “news” in the ECB announcement. Indeed, articles in the press confirm that market participants were expecting the ECB to reactivate its sovereign bond purchase program, and they were disappointed that it did not happen.³⁴ Therefore, the increase in sovereign spreads reflects mostly the market disappointment that a stronger measure, such as sovereign bond buying, was not announced to solve the euro-zone crisis.

The reaction of the sovereign spreads to cutting the ECB deposit rate to zero seems to confirm that sovereign bond markets in the euro zone were mostly driven by the market perception of the sovereign default risk. On July 5, 2012 the ECB reduced the deposit rate to zero but did not take any anti-crisis measures while the Italian and Spanish yields were very high. Therefore, this announcement (0 percent deposit at ECB in table 3) reduced significantly money-market spreads but increased Italian and Spanish sovereign spreads (+29 bp and +52 bp, respectively). The confirmation of this disappointment can be found in the press on that day.³⁵ The expected announcement came only three weeks later, with Mario Draghi’s “whatever it takes” speech (table 4), and made the Italian and Spanish spreads decrease significantly (−48 bp and −56 bp, respectively).

³⁴See, for instance, “US Stocks Fall as ECB Disappoints on Bond Buying,” *Wall Street Journal*, December 8, 2011, and “ECB Dampens Bond-Buying Hopes,” Reuters, December 8, 2011.

³⁵From “Spanish, Italian Bonds Hammered after Draghi Speech,” Reuters, July 5, 2012: “President Mario Draghi failed to deliver any hint that bolder monetary easing steps were on the way.” See also “European Stocks, Euro Drop after ECB Rate Cut,” Agence France-Presse, July 5, 2012: “The European Central Bank cut its key interest rate to a record low but did not unveil anti-crisis measures.”

The increase in sovereign spreads after the ECB failed to deliver strong anti-crisis measures expected by the markets highlights the challenges of unconventional monetary policies. Once the conventional monetary framework lost its primary role, it was much more difficult for the central bank to steer market expectations.³⁶

Overall, we show that the measures that diminished sovereign risk were the most successful in reducing government borrowing costs. First, sovereign bond purchases proved the most effective in countries confronted with high default probability. Second, bank covered bond purchases also diminished sovereign spreads in the context of the sovereign-bank risk interdependence.

5. Conclusion

The presence of sovereign risk is an important factor for the unconventional monetary policy implementation. The euro-zone sovereign debt crisis contributed to a dramatic increase in risk premia not only on sovereign debt markets but also on money markets and covered bond markets. The ECB's traditional tool, interest rate decrease, did not prevent the market borrowing costs from diverging across the member states. The ECB's unconventional monetary policies, however, contributed significantly to soothing financial tensions in the euro zone. We find that three-year LTROs and setting the ECB deposit rate to 0 percent, unlike the shorter-term liquidity measures, were effective in reducing money-market tensions. Furthermore, the measures aimed at diminishing the sovereign default risk proved the most effective in lowering longer-term risk premia. The ECB sovereign bond purchases (SMP and OMT) reduced significantly not only the sovereign spreads but also the bank covered bond spreads. Moreover, the ECB covered bond purchase programs diminished sovereign spreads, suggesting that the sovereign-bank feedback loop might amplify the effectiveness of the ECB's asset purchases in the euro zone.

³⁶ Another example of expectations management difficulty is Ben Bernanke's testimony before Congress on May 22, 2013 that markets interpreted as an announcement of early tapering of unconventional monetary policies. This message led to substantial turmoil in the financial markets despite its moderate character.

Table 5. Description of Dummies Used in the Regressions

<i>Unconventional Monetary Policies Dummies</i>	
Asset Purchases:	
SMP	Announcement of the longer-term sovereign bond purchase program (Securities Markets Programme, SMP). The ECB Governing Council decided to launch the SMP in the midst of the Greek crisis on Sunday, May 9, 2010. ^a A related press release was published the next day.
SMP2	Reactivation of SMP announced on Sunday, August 7, 2011, when the Spanish and Italian sovereign bond markets came under tension.
Draghi Speech	M. Draghi's speech of July 26, 2012 in which he promised to do "whatever it takes to save the euro." This speech triggered the expectations of ECB intervention in the sovereign markets.
OMT	Announcement of the ECB short-term sovereign bond purchase program (Outright Monetary Transactions, OMT) on September 6, 2012. No purchases were conducted within this program as of July 2014.
CBPP1 and CBPP2	Announcements of two ECB covered bond purchase programs (CBPP1 and CBPP2) on May 7, 2009, and on October 6, 2011. Both programs were designed to purchase covered bonds, i.e., securities issued by credit institutions backed by mortgages and public-sector debt.
Exceptional Liquidity Provisions:	
Three-Year LTRO An.	Announcement of the three-year longer-term refinancing operations (LTROs) on December 8, 2011. ECB banks were given the opportunity to obtain all liquidity needed for the period of three years at the interest rate fixed at the average rate of the main refinancing operations over the life of the operation.
Three-Year LTRO Op.	Dates of the two three-year LTROs conducted by the ECB. The first three-year LTRO was conducted on December 21, 2011 and the second on February 29, 2012.
FRFA	Announcements of the fixed-rate full-allotment (FRFA) procedure for different open-market operations: main refinancing operations (late on October 8, 2008), U.S. dollar funding (October 13, 2008), and LTROs (October 15, 2008). The reactivation of the FRFA procedure was announced on May 9, 2010.
0% Deposit	Announcement of the reduction to 0 percent of the interest rate on the ECB deposit facility.

(continued)

Table 5. (Continued)

Sovereign-Crisis-Related Dummies	
Sovereign Crisis	The sovereign crisis dummy captures the effects of the sovereign debt crisis in the euro zone. It is equal to 1 when a euro-zone member faces severe debt tensions: January 27, 2010–June 18, 2010; October 15, 2010–December 3, 2012; March 7, 2011–December 23, 2011; March 19, 2012–June 22, 2012. The crisis periods were defined using Google Insights,^b which graphically shows how often the search term “euro-zone sovereign debt crisis” is entered relative to the total search volume across various regions of the world. The horizontal axis of the main graph represents time, and the vertical axis represents how often a term is searched relative to the total number of searches. The results were cross-checked with euro-zone crisis timelines provided by Reuters, the <i>Wall Street Journal</i>, and the <i>Daily Telegraph</i>.
EF SF/ESM	Announcements concerning important extensions of two bailout facilities created by the euro-zone member states: the European Financial Stability Fund (EFSF) and the European Stability Mechanism (ESM). The EFSF was created on May 9, 2010 and authorized along with the ESM to intervene in the primary markets for sovereign debt on Saturday, March 12, 2011. On March 26, 2012, A. Merkel agreed to increase the overall ceiling for EFSF/ESM lending to €700 billion, up from €500 billion. On June 29, 2012, the euro-zone leaders allowed the euro zone’s rescue funds to recapitalize Spain’s enfeebled banks directly and to buy Italian sovereign bonds.
^aWhenever the announcement is made on weekends or after the market closure, we attribute a dummy equal to 1 on the following day. ^bSince the beginning of October 2012, Google Insights no longer exists. Some of its functions are included in Google Trends.	

References

- Abbassi, P., and T. Linzert. 2011. "The Effectiveness of Monetary Policy in Steering Money Market Rates during the Recent Financial Crisis." ECB Working Paper No. 1328.
- Angelini, P., A. Nobili, and C. Picillo. 2011. "The Interbank Market after August 2007: What Has Changed, and Why?" *Journal of Money, Credit and Banking* 43 (5): 923–58.
- Ayuso, J., and R. Repullo. 2003. "A Model of the Open Market Operations of the European Central Bank." *Economic Journal* 113 (490): 883–902.
- Beirne, J., L. Dalitz, J. Ejsing, M. Grothe, S. Manganelli, F. Monar, B. Sahel, M. Sušec, J. Tapking, and T. Vong. 2011. "The Impact of the Eurosystem's Covered Bond Purchase Programme on the Primary and Secondary Markets." ECB Occasional Paper No. 122.
- Bernanke, B. S. 2010. "The Economic Outlook and Monetary Policy." Speech given at the Federal Reserve Bank of Kansas City Economic Symposium, Jackson Hole, Wyoming, August 27.
- Brunetti, C., M. di Filippo, and J. H. Harris. 2011. "Effects of Central Bank Intervention on the Interbank Market during the Subprime Crisis." *Review of Financial Studies* 24 (6): 2053–83.
- Coeuré, B. 2012. "Challenges to the Single Monetary Policy and the ECB's Response." Speech given at the Institut d'études politiques, Paris, September 20.
- De Pooter, M., R. F. Martin, and S. Pruitt. 2012. "The Effects of Official Bond Market Intervention in Europe." Mimeo, Board of Governors of the Federal Reserve System.
- Draghi, M. 2012. "Verbatim of the Remarks Made by Mario Draghi." Speech given at the Global Investment Conference, London, July 26.
- Eggertsson, G. B., and M. Woodford. 2003. "The Zero Bound on Interest Rates and Optimal Monetary Policy." *Brookings Papers on Economic Activity* 34 (1): 139–235.
- Eser, F., and B. Schwaab. 2013. "Assessing Asset Purchases within the ECB's Securities Markets Programme." ECB Working Paper No. 1587.
- European Central Bank. 2008. "Covered Bonds in the EU Financial System." (December).

- European Covered Bond Council. 2010. *European Covered Bond Fact Book*. (September). European Mortgage Federation/European Covered Bond Council.
- Falagiarda, M., and S. Reitz. 2013. "Announcements of ECB Unconventional Programs: Implications for the Sovereign Risk of Italy." Working Paper No. 1866, Kiel Institute for the World Economy.
- Ghysels, E., J. Idier, S. Manganelli, and O. Vergote. 2013. "A High Frequency Assessment of the ECB Securities Markets Programme." CEPR Discussion Paper No. 9778.
- Hamilton, J. D., and J. C. Wu. 2012. "The Effectiveness of Alternative Monetary Policy Tools in a Zero Lower Bound Environment." *Journal of Money, Credit and Banking* 44 (s1): 3–46.
- Joyce, M. A. S., A. Lasasosa, I. Stevens, and M. Tong. 2011. "The Financial Market Impact of Quantitative Easing in the United Kingdom." *International Journal of Central Banking* 7 (3): 113–61.
- Krishnamurthy, A., and A. Vissing-Jorgensen. 2011. "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy." *Brookings Papers on Economic Activity* 43 (2, Fall): 215–87.
- Lenza, M., L. Reichlin, and H. Pill. 2010. "Monetary Policy in Exceptional Times." *Economic Policy* 25 (6): 295–339.
- MacKinlay, A. C. 1997. "Event Studies in Economics and Finance." *Journal of Economic Literature* 35 (1): 13–39.
- Pattipeilohy, C., J. Willem van den End, M. Tabbae, J. Frost, and J. de Haan. 2013. "Unconventional Monetary Policy of the ECB during the Financial Crisis: An Assessment and New Evidence." DNB Working Paper No. 381.
- Peersman, G. 2011. "Macroeconomic Effects of Unconventional Monetary Policy Measures in the Euro Area." ECB Working Paper No. 1397.
- Szczerbowicz, U. 2011. "Are Unconventional Monetary Policies Effective?" CELEG Working Paper No. 1107, Dipartimento di Economia e Finanza, LUISS Guido Carli.
- Taylor, J. B., and J. C. Williams. 2009. "A Black Swan in the Money Market." *American Economic Journal: Macroeconomics* 1 (1): 58–83.
- Ueda, K. 2012. "The Effectiveness of Non-Traditional Monetary Policy Measures: The Case of the Bank of Japan." *Japanese Economic Review* 63 (1): 1–22.

- . 2013. “The Response of Asset Prices to Abenomics: Is It a Case of Self-Fulfilling Expectations?” CIRJE Discussion Paper No. F-885, Faculty of Economics, University of Tokyo.
- Wright, J. H. 2012. “What Does Monetary Policy Do to Long-Term Interest Rates at the Zero Lower Bound?” *Economic Journal* 122 (564): F447–F466.
- Vayanos, D., and J.-L. Vila. 2009. “A Preferred-Habitat Model of the Term Structure of Interest Rates.” CEPR Discussion Paper No. 7547.

Liquidity Requirements: A Double-Edged Sword*

Philipp Johann König
DIW Berlin

This paper shows that bank liquidity regulation may be a “double-edged sword.” Under certain conditions, it may hamper, rather than strengthen, a bank’s resilience to financial stress. The reason is the existence of two opposing effects of liquidity regulation, a liquidity effect and a solvency effect. The liquidity effect arises because a bank mitigates its risk of illiquidity when it increases its liquidity buffer. The solvency effect arises because a larger liquidity buffer reduces the bank’s returns and may therefore raise its insolvency risk. Liquidity regulation is effective in reducing a bank’s overall default risk only if the former effect dominates the latter. The paper derives conditions under which this is the case and discusses the resulting relationship between capital and liquidity regulation.

JEL Codes: G21, G28.

1. Introduction

The recent financial crisis was associated with severe liquidity problems in the banking sector. This experience has led regulators to put greater emphasis on issues of bank liquidity risk than was the case before the crisis. In particular, in the course of the latest reform of the Basel Accords on banking regulation (Basel III), new liquidity standards have been proposed in order to strengthen liquidity risk measurement and supervision. Two regulatory standards on bank

*I very much thank Rafael Repullo, the editor, and an anonymous referee for helpful comments and suggestions. I am also indebted to Frank Heinemann, Ulrike Neyer, Tijmen Däniels, Alexander Meyer-Gohde, Kartik Anand, David Pothier, Helmut Lütkepohl, and Marcel Fratzscher for comments, suggestions, and discussions. Author contact: DIW Berlin, Department of Macroeconomics, Mohrenstraße 58, 10117 Berlin, Germany; e-mail: pkoenig@diw.de.

liquidity have been introduced which shall complement and institutionalize the Basel Committee's guidelines for liquidity risk management.¹ These newly introduced liquidity standards are the "net stable funding ratio" (NSFR) and the "liquidity coverage ratio" (LCR). The respective objectives of these measures differ but complement each other. The LCR is designed to enhance banking institutions' resilience to short-term funding stress. It shall ensure that banks hold sufficiently liquid assets to meet outflows of short-term liabilities under a stress scenario lasting for around one month.² The NSFR shall promote the funding stability of banks over a longer period (i.e., up to one year). It requires them to fund their assets using sufficiently stable sources, thereby limiting banks' exposure to short-term wholesale funding.³

The present paper uses the well-known banking model by Rochet and Vives (2004) in order to provide some analytical results on the functioning of short-term liquidity buffers (like the LCR). It discusses whether, and under which circumstances, they actually serve regulators' microprudential objective and improve an individual bank's resilience in periods of financial stress.

By refining a result of Rochet and Vives (2004, proposition 3), it is demonstrated that liquidity requirements strengthen a bank's resilience (as measured by its probability of default) if and only if the likelihood of a default is sufficiently low. This result is further translated into a threshold level on capital requirements. Essentially, liquidity requirements are effective whenever the bank satisfies a certain minimum capital requirement. This allows to derive conclusions concerning the interaction between a bank's liquidity and solvency.

In general, the reason why liquidity requirements may only work effectively under particular circumstances is the existence of two opposing effects, a *liquidity effect* and a *solvency effect*. The liquidity effect arises because it is costly for the bank to use an illiquid asset as a cushion against sudden funding roll-offs. Converting such an asset into cash at short notice involves a loss, and keeping a larger fraction of the balance sheet in the form of liquid assets thus allows the bank to withstand larger funding drains without engaging in costly fire

¹See Basel Committee on Banking Supervision (2008).

²See Basel Committee on Banking Supervision (2013).

³See Basel Committee on Banking Supervision (2014).

sales. This in turn decreases the bank's risk of default due to illiquidity. The solvency effect occurs because liquid assets earn rather low returns on average and thus fail to generate the returns which the bank needs to service its interest-bearing deposit liabilities. Given that the bank manages to roll over its debt, the lower returns on liquid assets have to be compensated by sufficiently higher returns on more profitable (but less liquid) assets in order for the bank to remain solvent. This increases its risk of insolvency. Moreover, when creditors perceive the higher insolvency risk, their incentive to withdraw their funds may increase, thereby spurring the illiquidity of the bank. The existence of these competing effects implies that the use of liquidity requirements is a "double-edged sword" and that it is a priori not clear whether they make the bank more resilient to financial stress.

In their original analysis, Rochet and Vives (2004) neglect the solvency effect and focus only on the liquidity effect. However, their analysis is based on an implicitly assumed adjustment of the bank's balance sheet which, as will be explained below, happens to be mistaken. In what follows, I solve Rochet and Vives's model by making the underlying balance sheet adjustment explicit, thus taking insolvency risk into account.

The remainder of the paper is organized as follows. Section 2 provides a condensed description of Rochet and Vives's banking model.⁴ Section 3 contains the relevant comparative statics exercises and a discussion of liquidity and solvency effects. Section 4 considers how balance sheet levels adjust under different assumptions concerning the determination of the bank's balance sheet scale. Section 5 contains two extensions of the model in order to study the robustness of the baseline results. Section 6 concludes. All proofs and most of the mathematical derivations are relegated to the appendix.

2. The Rochet-Vives Banking Model

The model studies a bank that operates at three dates, labeled 0, 1, and 2. At date 0, the bank has equity E_0 and takes in wholesale deposits D_0 . The deposit contract promises a repayment $D > D_0$

⁴For a full derivation of the model, the interested reader is referred to the original paper.

independent of whether the funds are withdrawn at date 1 or 2. The bank uses amount I to fund a risky asset and holds amount M in cash. Its balance sheet at date 0 reads

$$I + M = D_0 + E_0. \quad (1)$$

Cash holdings are non-interest bearing but they are perfectly liquid in the sense that one unit of cash can be used to cover one unit of liabilities at date 1 or 2.⁵ The per-unit return on the risky asset is a normally distributed random variable $\tilde{R}$ with mean $\bar{R} > 1$ and precision (inverse variance) α . The asset pays out the realized return R at date 2. At date 1, the bank can sell the asset on a secondary market against a fire-sale price $R/(1 + \lambda)$ to obtain additional cash. The fire-sale discount rate $\lambda > 0$ renders the asset illiquid, which implies that the failure of the bank may depend on the fraction of deposits withdrawn early at date 1.

The bank fails at date 2 whenever

$$R < R_s + \frac{\lambda \max \{xD - M, 0\}}{I}, \quad (2)$$

where

$$R_s \equiv \frac{D - M}{I} \quad (3)$$

is henceforth called the bank's solvency point and $x \in [0, 1]$ denotes the proportion of deposits withdrawn. To gain the intuition behind equation (2), note that in case $xD < M$, the bank can cover all withdrawals out of its cash holdings and does not need to resort to fire sales of the illiquid asset on the secondary market. It can only fail due to *insolvency* at date 2, i.e., if $R < R_s$. In contrast, if $xD > M$, its cash holdings are not large enough and the bank has to fire-sell assets to cover early withdrawals. Compared with a situation where it would not have sold assets, the bank loses $\lambda R/(1 + \lambda)$ per unit sold. Total losses rise with x and may deplete the bank's resources to such

⁵The assumption that the bank holds cash is not material for the results. Alternatively, the bank may hold other types of liquid assets. The only important assumption is that the returns on the liquid assets are lower than the interest rate on deposits.

an extent that it becomes unable to cover the remaining liabilities $(1-x)D$ at date 2 and defaults. This explains why the failure point after fire sales exceeds the solvency point by a factor proportional to the fire-sale discount rate. If the bank fails at date 2 even though it is solvent ($R > R_s$), it is said to fail due to *illiquidity*.⁶

It is further assumed that the management of deposits and the withdrawal decisions are made by fund managers on behalf of the original depositors. The bank has contracted with a large number of ex ante identical fund managers. Their total size is normalized to the unit interval, and each fund manager administers an amount D_0 .⁷

Fund managers receive a remuneration B . However, when they withdraw early at date 1, they incur an additional cost C . Moreover, in case of bank default, a fund manager who decided to roll over is fired and receives a zero payoff.⁸ The optimal decision of fund managers is to withdraw at date 1 whenever the expected payoff from withdrawing exceeds the expected payoff from rolling over. Formally, a typical fund manager $i \in [0, 1]$ withdraws if and only if

$$P_i > \gamma \equiv \frac{C}{B},$$

where P_i denotes the subjective probability that fund manager i attaches to the failure of the bank.

⁶The bank defaults already at date 1 if withdrawals fully exhaust its cash holdings and the liquidity obtained from fire-selling all its assets. This case will be taken up in section 5 below when the deposit interest rate is endogenized. However, for the determination of the equilibrium default probability, the relevant case to study is default at date 2 because it occurs for a larger range of return realizations than default at date 1; see Rochet and Vives (2004, p. 1126).

⁷The assumption that foreclosure decisions are carried out by fund managers instead of by fund owners themselves may reflect the fact that a major part of the short-term funding base of banks, especially in the United States, is managed by money-market mutual funds or collective investment funds. Experience during the recent crisis has shown that especially this part of bank funding is more volatile and evaporates quicker than the retail and customer deposit base. The assumption seems therefore reasonable for studying the short-term funding liquidity situation of a large bank that is subject to liquidity regulation.

⁸For a motivation of the modeling of fund managers' payoffs, see the explanations in Rochet and Vives (2004), who refer to the empirical studies of Chevalier and Ellison (1997, 1999).

Fund managers do not observe R but instead receive an idiosyncratic noisy signal about R at date 1.⁹ The signal takes the form

$$s_i = R + \varepsilon_i,$$

where ε_i is i.i.d. normally distributed with zero mean and precision β . Fund managers then calculate their (posterior) belief P_i which equals the probability (conditional on s_i) that R is below the upper bound identified in condition (2).

The model satisfies the standard requirements needed to apply global game methods for the derivation of a unique equilibrium.¹⁰ The unique equilibrium in this class of models is a symmetric monotone (or threshold) equilibrium. That is, there exists a threshold signal t^* such that a typical fund manager withdraws whenever $s_i < t^*$ and rolls over otherwise, as well as a critical threshold R^* such that the bank defaults whenever $R < R^*$ and survives otherwise. By the law of large numbers, the proportion of deposits withdrawn at date 1 can then be computed as the probability (conditional on R) that a single signal s_i is below t^* . The equilibrium is unique provided that the private signals are sufficiently precise. In the Rochet-Vives model, the equilibrium thresholds (t^*, R^*) solve the pair of equations

$$\Phi\left(\sqrt{\alpha + \beta}R^* - \frac{\alpha\bar{R} + \beta t^*}{\sqrt{\alpha + \beta}}\right) = \frac{C}{B} \quad (4)$$

and

$$R^* = R_s + \frac{\lambda \max\{\Phi(\sqrt{\beta}(t^* - R^*))D - M, 0\}}{I}, \quad (5)$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function.¹¹

⁹Under common knowledge of R , the model would exhibit multiple self-fulfilling equilibria for $R \in [R_s, (1 + \lambda)R_s)$. This outcome would, for example, be equivalent to the outcome in the well-known model by Diamond and Dybvig (1983). For a discussion of the influence of the assumption of common knowledge of R on the equilibrium multiplicity, see Morris and Shin (2001).

¹⁰See Morris and Shin (2003) or Heinemann (2005) for overviews of the global game methodology.

¹¹See Rochet and Vives (2004, p. 1129).

Given the bank's failure point R^* , equation (4) is derived from the equilibrium requirement that a fund manager who uses a threshold strategy around t^* must be indifferent between rolling over and withdrawing if he observes a signal s_i exactly equal to the critical threshold t^* . Similarly, for a given t^* (determining the proportion of withdrawals x), equation (5) is derived from the requirement that at $R = R^*$, equation (2) must hold with equality. Furthermore, as shown by Rochet and Vives (2004, proposition 1), the equilibrium is unique provided that the precision of fund managers' signals is sufficiently large, i.e., if

$$\beta \geq \beta_0 \equiv \frac{\alpha^2}{2\pi} \left(\frac{\lambda D}{I} \right)^2. \quad (6)$$

Inequality (6) is the analogue to the standard uniqueness condition in global games with normally distributed fundamental and noise variables.¹² Whenever it fails to hold, equilibrium multiplicity may reemerge. Essentially, in this case, the weight on the publicly known expected return $\bar{R}$ in the formation of fund managers' beliefs is sufficiently large so that public information dominates private information and the outcome of the game approaches the outcome of a game where R is common knowledge and multiple self-fulfilling equilibria exist.

3. Analysis and Discussion of the Model

3.1 Comparative Statics

The ex ante default probability, given by

$$\Pr(\tilde{R} \leq R^*) = \Phi(\sqrt{\alpha}(R^* - \bar{R})),$$

is used as a measure of the bank's resilience against financial stress. Since $\Phi(\cdot)$ is a strictly monotone function, when studying the effects of regulatory ratios on the ex ante default risk, it suffices to study their effects on the failure point R^* . As the Basel Accords identify

¹²See Hellwig (2002) or Morris and Shin (2003).

the strengthening of banks' resilience as a major objective of micro-prudential regulation, the comparative statics properties of R^* may indicate to which extent regulatory requirements may or may not meet this objective.

Following Rochet and Vives, I focus on the effects of two regulatory ratios—the *liquidity ratio*, given by $m \equiv M/D$, and the *capital ratio*, given by $e \equiv E_0/I$. These ratios can be viewed as approximations to the Basel Accord's newly introduced LCR and the capital adequacy ratio.¹³ Differently from Rochet and Vives, in the comparative statics below, the ratio of deposits to risky investments, $d \equiv D_0/I$, is treated as endogenous. To see its dependence on e , m , and the interest rate on deposits, $r \equiv D/D_0 - 1$, observe that the balance sheet constraint (1) can be rewritten as

$$d = \frac{1 - e}{1 - (1 + r)m}. \quad (7)$$

Moreover, the solvency point can be expressed as

$$R_s = d(1 + r)(1 - m). \quad (8)$$

By substituting equation (7) into (8), the solvency point can be expressed as a function of m , e , and r ,

$$R_s(m, e, r) = \frac{(1 - e)(1 + r)(1 - m)}{1 - (1 + r)m}. \quad (9)$$

In a similar way, equation (5) can be written as

$$R^* = R_s(m, e, r) + \frac{\lambda(1 - e)(1 + r) \max \{ \Phi(\sqrt{\beta}(t^* - R^*)) - m, 0 \}}{1 - (1 + r)m}. \quad (10)$$

Whenever $x > m$, the failure point R^* exceeds the solvency point R_s . In this case, R^* is given by the solution to the implicit function,

¹³Rochet and Vives label the ratio E_0/I as *solvency ratio*, which I found somewhat misleading, as this ratio affects not only the insolvency but also the illiquidity risk of the bank.

$$\begin{aligned}
\psi(R^*, m, e, r) \equiv & (1 - m(1 + r))R^* - (1 - e)(1 + r)(1 - m) \\
& + \lambda(1 - e)(1 + r)m \\
& - \lambda(1 - e)(1 + r)\Phi\left(\frac{\alpha}{\sqrt{\beta}}(R^* - \bar{R})\right. \\
& \left. - \sqrt{\frac{\alpha + \beta}{\beta}}\Phi^{-1}(\gamma)\right) = 0. \tag{11}
\end{aligned}$$

This equation is derived by solving equation (4) for t^* and then substituting the result together with the expression for R_s from equation (9) into (10).

The comparative static properties of R^* are summarized in the following proposition.¹⁴

PROPOSITION 1. *The failure point R^* and the likelihood of bank default strictly decrease in the capital ratio e and increase in the interest rate on deposits r .*

If $R^* = R_s$, the failure point and the likelihood of default increase in the liquidity ratio m .

Otherwise, if $R^* > R_s$, they strictly decrease in m if and only if

$$R^* < (1 - e)(1 + \lambda). \tag{12}$$

3.2 Liquidity and Solvency Effects

The description of the model in section 2 proceeded closely along the lines of Rochet and Vives (2004). But the approach to the comparative statics and the conclusion concerning the effects of the liquidity ratio on the failure point R^* differ. This is due to a different treatment of the bank's balance sheet constraint in the derivation of the failure point. While proposition 1 confirms Rochet and Vives's results that the failure point always decreases in the equity ratio e and increases in the interest rate r (see their proposition 3), the effect of the liquidity ratio m depends on the underlying model parameters. This result stands in contrast to Rochet and Vives (2004,

¹⁴As is common in the literature on global games, the comparative statics are conducted under the assumption that the uniqueness condition (6) is satisfied.

proposition 3), as they claim that an increase in the liquidity ratio unambiguously lowers the failure point and the default probability. The reason for this difference is that Rochet and Vives first use the balance sheet constraint to express R_s as a function of e and d , but then neglect the dependency of d on m when conducting the comparative statics with respect to m (recall equation (7)). Accordingly, in their analysis, for a given e , a change in m does not lead to an adjustment of R_s , and their assessment of the effects of the liquidity ratio on bank solvency becomes incorrect.¹⁵

By treating R_s as independent from m , Rochet and Vives suppress what I will henceforth call the *solvency effect*. To better appreciate the intuition behind it, note that, from equations (4) and (5), the failure point R^* and the switching signal t^* depend positively on the solvency point R_s .¹⁶ Moreover, the solvency point R_s is strictly increasing in the liquidity ratio m : By definition, a higher liquidity ratio is equivalent to a higher level of non-interest-bearing assets relative to interest-bearing liabilities. Since the balance sheet constraint (1) holds, the bank needs to achieve higher net returns to service its deposit liabilities, i.e., without generating a higher rate of return on its profitable asset, the risk of becoming insolvent rises.

Whenever $R^* = R_s$, only the solvency effect is present and a higher liquidity ratio m cannot be used to effectively lower the bank's default risk. However, if $R^* > R_s$, an otherwise solvent bank may also fail due to illiquidity. In this case, a change in the liquidity ratio m is associated with a second effect which works in the opposite direction and which may be called the *liquidity effect*. Since illiquidity is essentially governed by the depletion of the bank's assets due to fire sales at date 1, a higher liquidity ratio m reduces the

¹⁵Keeping R_s constant upon a marginal change in m requires to adjust d by $dd = (d/(1-m))dm$ (see equation (8)). But from the balance sheet constraint (7), this implies $de = (rd/(1-m))dm \neq 0$. Hence, it is impossible to adjust m for fixed e and R_s without violating the balance sheet identity.

¹⁶One may view this dependency as a formal restatement of Charles Goodhart's notion that illiquidity implies "at least a whiff of suspicion of insolvency" (1999, p. 346). If the insolvency risk increases (i.e., R_s increases), fund managers become more "suspicious" and attach a higher likelihood to the event that the bank fails. This increases the switching point t^* and therefore the failure point R^* .

depletion of profitable assets and thereby lowers the bank's exposure to illiquidity. Hence, the liquidity effect decreases the failure point R^* . Consequently, the use of liquidity requirements as a prudential measure to strengthen the bank's resilience crucially depends on whether the liquidity or the solvency effect dominates.

Condition (12) shows that the liquidity effect dominates the solvency effect whenever the failure point, i.e., the bank's default risk, is sufficiently low. However, since R^* is an endogenous object which itself depends on exogenous parameters—in particular on e and m —it is not immediately apparent how changes in these parameters affect condition (12). To derive a condition for the dominance of either liquidity or solvency effect purely in terms of exogenous parameters, equation (11) can be evaluated at $R^* = (1 + \lambda)(1 - e)$,

$$\psi((1 + \lambda)(1 - e), m, e, r) = 0,$$

thus yielding the particular combination of exogenous parameters for which liquidity and solvency effect cancel each other out completely so that R^* does not respond to m anymore. By solving the above equation for r , a threshold value for the interest rate can be derived,

$$r^m(e) = \frac{\lambda(1 - \Phi(\omega(e)))}{1 + \lambda\Phi(\omega(e))} \quad \text{with}$$

$$\omega(e) \equiv \frac{\alpha}{\sqrt{\beta}}((1 + \lambda)(1 - e) - \bar{R}) - \sqrt{\frac{\alpha + \beta}{\beta}}\Phi^{-1}(\gamma), \quad (13)$$

such that $R^* = (1 + \lambda)(1 - e)$ if and only if $r = r^m(e)$. As a reduction of r decreases the failure point R^* , it follows that the liquidity effect dominates the solvency effect if and only if the interest rate on deposit liabilities is sufficiently low, i.e., $r < r^m(e)$. To appreciate the intuition behind this result, suppose that $r = 0$. The face value of the bank's liabilities equals the amount of initially obtained deposits. The bank does not need to make interest-bearing investments to pay back its deposits; rather, it may simply store them in cash. Moreover, from equation (9) follows that R_s ceases to depend on m . Thus, only the liquidity effect is present and the failure point R^* strictly decreases in m . But with positive interest rates, $r > 0$, the face value of liabilities exceeds the value of initial deposits. Per unit of deposits held, the bank needs to obtain positive net returns from

its investments to become able to service its debt. This, however, is not possible by resorting to non-interest-bearing cash holdings. A fortiori, if we keep increasing r , the ability of the bank to cover withdrawals out of its cash holdings weakens while the solvency point R_s increases more and more in m . With r being above the critical value r^m , the solvency effect becomes too strong and eventually dominates the liquidity effect.

Another way to look at liquidity and solvency effect and to spell out the implied relationship between liquidity and capital regulation more clearly is by inverting the threshold rate $r^m(e)$ to obtain a critical capital ratio $e^m(r)$. As $r^m(e)$ strictly increases in e , the liquidity effect is dominant for $e > e^m$. In this case, liquidity and capital regulation are substitutes: A lower capital ratio can be compensated by a higher liquidity ratio without changing the risk level R^* . However, if the capital ratio falls short of e^m , this relationship suddenly reverses. While increases in the capital ratio still produce a lower default risk, increases in the liquidity ratio have the opposite effect. Consequently, in order to keep the bank's risk level fixed, increases in m have to be complemented by a higher level of e .

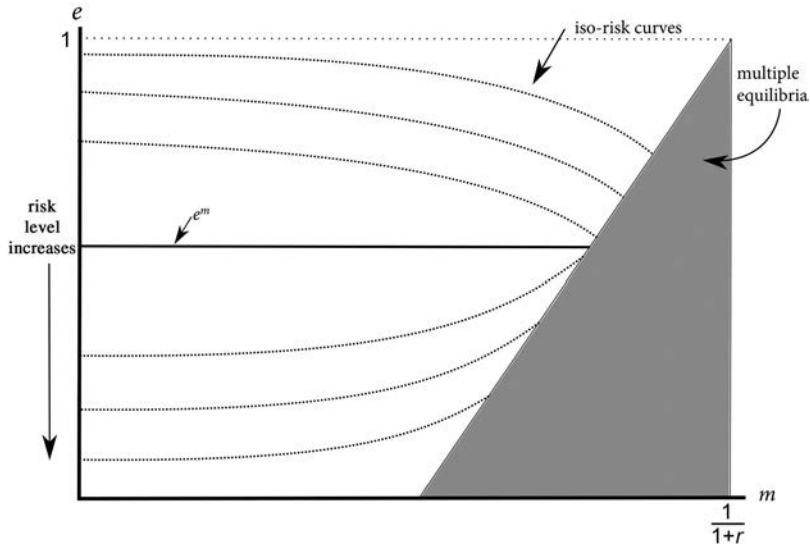
This relationship is also illustrated in figure 1, which plots iso-risk curves in m - e space.¹⁷ Each curve shows the combinations of m and e needed to keep the risk level fixed. The risk level is lowest at the top curve and increases from top to bottom. For a given risk level—say, R_k^* —the iso-risk curve $e = e(m; R_k^*)$ is implicitly defined through equation (11).¹⁸ Along the horizontal line at $e = e^m$, variations in m exert no effect on the risk level.¹⁹ Above this line, the liquidity effect is dominant and iso-risk curves are downward sloping, thus reflecting the substitutability between liquidity and capital regulation. Conversely, below the line, iso-risk curves are upward sloping, reflecting the dominating solvency effect and the fact that increases in m need to be accompanied by increases in e to keep the risk level

¹⁷The domain of e is taken to be the unit interval and the domain of m is the interval $[0, 1/(1+r)]$: Firstly, neither equity nor cash holdings can become negative, so e and m are bounded below by zero; secondly, given an ex ante expected return $\bar{R} > 1$, the bank would invest at least its equity into the risky asset, implying that $e \leq 1$ and $m \leq 1/(1+r)$.

¹⁸The derivation of iso-risk curves is provided in the appendix.

¹⁹ e^m is by definition the value of e where the derivative of R^* with respect to m vanishes. This implies that e^m is itself independent of m and the iso-risk curve at $e = e^m$ is a horizontal line.

Figure 1. Iso-Risk Curves in m - e Space



fixed. This shows that liquidity requirements can be a double-edged sword with respect to their effects on default risk. Moreover, the existence of structural differences across banks implies that liquidity regulation may have to be bank-specific as long as homogeneous capital standards are applied. That is, in order to stabilize banks' risk levels following an adverse shock, some banks would have to be advised to lower their liquidity ratios, while other banks would need to raise them. This may impede the use of general liquidity rules applying uniformly to all banks.

Finally, the shaded area in figure 1 shows combinations of e and m such that for given parameters β , α , r , and λ , multiple equilibria emerge. The boundary line of this area can be calculated by using equation (7), substituting $d(1+r) = D/I$ into condition (6), thus yielding (with equality)

$$\beta = \frac{\alpha^2}{2\pi} \left(\frac{\lambda(1+r)(1-e)}{1 - (1+r)m} \right)^2 \quad (14)$$

and solving the latter for e in terms of m .

The point where any iso-risk curve touches the boundary line constitutes a bifurcation point from which essentially three equilibria emerge.²⁰ Since each combination of e and m is then associated with three different risk levels, iso-risk curves are only drawn until they touch the boundary.

For given parameter values, a higher capital ratio slackens condition (14) and implies that a unique equilibrium obtains for a larger range of values for β . In contrast, a higher liquidity ratio m has the opposite effect. In particular, if for given β the liquidity ratio is raised above a certain point, the unique equilibrium breaks down and multiple self-fulfilling equilibria occur. This can be seen by going back to condition (6) and noting that the condition becomes tighter when $\lambda D/I$ increases. A higher D/I raises the bank's dependency on short-term deposits and a higher λ increases potential fire-sale losses. Hence, the larger $\lambda D/I$, the higher the bank's vulnerability to a coordination failure and to self-fulfilling bank runs. In order to ensure uniqueness of the global game solution, this effect has to be balanced by more precise private signals of fund managers. As D/I depends negatively on e and positively on m , the boundary line is upward sloping.²¹

4. Balance Sheet Adjustment

The previous section showed that a bank's risk level is determined by the two ratios e and m . Yet, focusing only on ratios consisting of balance sheet components leaves open how their levels adjust to variations in these ratios. For example, when changing e and m while moving along an iso-risk curve, the underlying balance sheet may also change. In order to analyze these balance sheet adjustments in greater detail, I assume that e and m are completely determined by the regulatory authorities. This may reflect a situation where the

²⁰This result was first explored in greater detail by Hellwig (2002). See also the discussion in Morris and Shin (2003).

²¹With respect to the impact of the liquidity ratio, it can be seen from equation (7) that D/I strictly increases in m . Intuitively, keeping D fixed, investments have to be lowered to raise cash holdings in order to increase m (thus D/I increases); alternatively, if I is fixed, m can be increased by raising new deposits and storing them in cash (again increasing D/I).

bank would choose too-low levels of e and m in the absence of regulatory intervention. For example, preferential tax treatment of debt due to the deductibility of interest payments from corporate taxes (debt tax shield), or the desire to avoid incurring large (opportunity) cost by holding cash may induce banks to minimize e and m . This, however, may conflict with the regulator's objectives of financial stability and more resilient banks. The regulator therefore enforces certain minimum levels of e and m . This way, the definitions of the two ratios and the balance sheet constraint provide three equations to determine the four balance sheet components E_0 , D_0 , I , and M . Thus, in order to close the model, one further assumption is needed. To this end, I distinguish between three border cases where either E_0 , D_0 , or I is fixed exogenously and then consider how the balance sheet adjusts to regulatory changes in e or m . The following three equations govern the balance sheet adjustment:

$$\frac{dM}{M} - \frac{dD_0}{D_0} = \frac{dm}{m} \quad (15)$$

$$\frac{dE_0}{E_0} - \frac{dI}{I} = \frac{de}{e} \quad (16)$$

$$dM - dD_0 - dE_0 + dI = 0. \quad (17)$$

Exogenous Equity Level E_0 . A fixed equity level may approximate a situation where the bank finds it overly costly to raise additional outside equity. Alternatively, as equity is a variable that needs relatively more time to be adjusted than, for example, debt, one may also think of this scenario as the short run where the equity level cannot be changed. In this case, the maximal investment scale is determined by the capital ratio. Consequently, if the bank always chooses the maximal investment scale, an inverse relationship between capital ratio e and investment I emerges. If the regulator raises, say, the capital ratio, the bank adjusts by scaling down its investments by $-Ie^{-1}de$. In addition, as the bank avoids holding too much idle cash and keeps the liquidity ratio constant at the regulatory minimum, the total balance sheet size shrinks. Given that I decreases while E_0 is constant, the lower level of investments requires a lower level of deposits and, to keep m constant, the level of cash holdings M can be lowered as well.

If the regulator enforces a higher liquidity ratio while keeping the capital ratio constant, I cannot be adjusted, and it follows from equations (15) and (17) that $dD_0 = dM = D \times dm / (1 - (1+r)m) > 0$. The bank's balance sheet lengthens and its indebtedness rises, but any new deposits are stored entirely in cash. Leaving a costly resource lying in idle cash may look like a bad decision from the point of view of portfolio management. But it may be the right decision from the regulator's point of view if the additional cash allows to reduce the bank's default risk by shielding it against fire-sale losses.

This example allows to discuss how regulations linking capital buffers with business or financial cycles in a countercyclical way (capital requirements which are loosened during downswings and tightened during upswings) may give rise to a trade-off between stabilizing an individual bank's default risk and the level of investment. On the one hand, if the capital ratio is lowered during business-cycle downswings, a bank can increase the level of its investments. This may help to stabilize the economy in the short run. On the other hand, a lower capital ratio increases the bank's default risk. What's more, whenever the bank's default risk negatively correlates with the business cycle—for example, due to fluctuations in the expected rate of return $\bar{R}$ —countercyclical capital buffers amplify risk increases in the banking sector during business-cycle fluctuations. Is it possible to overcome this dilemma by resorting to additional liquidity regulation? It requires the regulator to know whether a bank is located on an iso-risk curve above or below the e^m line and how e^m itself reacts to the cycle. For example, the critical level e^m depends negatively on the expected return $\bar{R}$, which may positively correlate with business-cycle swings. A downswing therefore increases the area in figure 1 where the solvency effect dominates by shifting up the e^m line. If, following a decline in $\bar{R}$, the bank will be located on an iso-risk curve below the new e^m line, the regulator must advise a lower liquidity ratio to counter the higher risk level and vice versa if the bank will be above the new line. Hence, it may be possible to counter adverse effects of countercyclical capital buffers by adequate additional liquidity regulation. But for liquidity regulation to be indeed adequate, the regulator must know a bank's precise position in e - m space relative to the prevailing e^m line.

Exogenous Investment Scale *I*. Next, consider the case where the investment scale, rather than the level of equity, is exogenously given. One may think here of a smaller bank with a local monopoly in supplying loans or, alternatively, of a bank focusing on a particular clientele. The capital ratio now determines the necessary minimum amount of equity that the bank has to raise. Given that the bank would economize on the level of equity in order to maximize shareholders' rate of return, a positive relationship between the equity level and the capital ratio emerges.²² If the bank tries also to economize on its cash balances, an increase in the capital ratio will be accompanied by a lower level of deposits and smaller cash holdings in order to keep the liquidity ratio constant. Adjustments in the liquidity ratio proceed in the same way as in the above case with exogenous E_0 : investment and equity levels are not affected, and a higher liquidity ratio is achieved by taking in new deposits which are entirely kept in cash.

What happens if the bank faces new investment opportunities—i.e., $dI > 0$ —but has to maintain given levels of e and m ? Firstly, it has to fund a fraction e of new investments by issuance of equity and the remaining fraction $(1 - e)$ by new deposits. But secondly, as these new deposits would lower the liquidity ratio m , it has to raise an even larger amount of funds in order to fill up its cash buffer,

$$dD_0 = (1 - e)dI / (1 - (1 + r)m) > (1 - e)dI.$$

This case may be used to shed some light on the situation in the banking sector in the absence of liquidity regulation. Prior to the crisis of 2008/2009, banking regulation almost exclusively focused on capital regulation. As observed by Goodhart (2009), banks shifted into less liquid assets and liquidity ratios deteriorated in the run-up to the crisis. In the model, under regulation focusing exclusively on capital, exogenous shifts in investment opportunities may lead to a deterioration of the liquidity ratio. If capital ratios are sufficiently high—i.e., above e^m —worsening liquidity ratios entail higher default risk by rendering banks more vulnerable to a sudden dry-up in funding markets. This counteracts, to some extent, the efforts made by

²²In practice, the effectiveness of capital regulation in the short run may be impeded by the fact that adjustments of equity levels are more time-consuming and costly than adjustments of debt levels.

implementing high capital ratios. Put differently, in order to steer the risk level of banks and to weaken the link between variations in banks' portfolios and changes in their default risk, the regulator must pin down both capital and liquidity ratios.

Exogenous Debt Level D_0 . Finally, consider the case where the deposit level is exogenously given. This case does not fully square with the interpretation of the model given above, where D_0 was interpreted as wholesale funds administered by fund managers. However, suppose for a moment that D_0 consists of uninsured retail funds that the bank cannot increase easily of its own accord. In this case, the investment level I is determined by the interplay between the regulatory ratios e and m . Using the balance sheet constraint given in equation (7), changes in the investment level can be computed as

$$dI = \frac{-(1+r)D_0}{1-e}dm + \frac{(1-m(1+r))D_0}{(1-e)^2}de.$$

A higher capital ratio e tends to induce an increase in I since the bank could use the newly issued equity to fund additional investments. Conversely, a higher liquidity ratio would lead to a reduction in the investment scale because, as $D = D_0(1+r)$ is fixed, the bank would need to swap investments for cash holdings to meet regulatory requirements. Then, moving from left to right along an iso-risk curve located above e^m , the investment level declines as the bank successively reallocates its funding into cash holdings. Below e^m , however, movements along the iso-risk curve are associated with an ambiguous effect since an increase in cash holdings is accompanied by additional issuances of equity and there exists a counterbalancing effect on the investment scale.

5. Extensions

This section considers two extensions of the model in order to examine the robustness of the liquidity and solvency effect with respect to the assumptions of exogenously given interest rate r and fire-sale rate λ . Treating these as parametric has facilitated the analysis of the model so far. Yet, it may be important to understand to which extent liquidity and solvency effects are modified once these variables are treated endogenously.

5.1 *Endogenous Interest Rate*

The baseline model in section 2 uses the rather extreme assumption that the interest rate does not depend on the bank's default risk and therefore does not react to changes in liquidity and capital ratios. To examine the importance of this assumption, I now turn to the opposite extreme and consider a variant of the model where the interest rate embodies a risk premium and perfectly reflects the bank's default risk. A straightforward way to capture this dependency is to require that, in expected terms, risk-neutral creditors have to be at least indifferent between investing with the bank or putting their funds into a risk-free two-period asset with exogenous interest rate $\bar{r}$. Thus, the fund managers take care not only of the interim withdrawal decision but also of the initial investment decision. Under the assumption of a perfectly competitive money market, parity in expected terms between the risk-free rate and the interest rate on deposits will hold in equilibrium.

In setting up the relevant interest parity condition, it becomes important to distinguish between default at date 1 (early default) and default at date 2, since the bank's liquidation value (and thus the amount paid out per creditor in case of default) differs between these cases.²³

The bank defaults early at date 1 if it cannot obtain enough liquidity (cash plus secondary market liquidity from fire-selling assets) to meet withdrawals. This arises whenever

$$R < (1 + \lambda) \frac{x D - M}{I}. \quad (18)$$

Suppose that withdrawing creditors face a sequential service constraint, i.e., the bank pays creditors until its resources are exhausted. In case the bank defaults early at date 1, its liquid resources are insufficient to meet total withdrawals, and a withdrawing creditor

²³In general, the creditor's cost of delegating the management of the funds to the fund manager, i.e., $B - C$ in the case of early withdrawal and B in the case of successful rollover, would also need to be considered. Since only the ratio $C/B \equiv \gamma$ matters for the analysis, I can assume that the costs of managing D_0 units of cash are small, so that neither B nor C affect the initial investment decision, while the ratio γ is still well defined.

receives his claim only with a certain probability.²⁴ Creditors whose claims are rolled over get nothing. However, if the bank defaults at date 2, withdrawals at date 1 are met in full, while those creditors who roll over receive a pro rata share of the remaining liquidation value of the bank.²⁵

This modified version of the model then consists of three endogenous variables, the interest rate on deposits and the date 1 and date 2 failure points. The model can be solved by first determining the two default points from equations (11) and (18). The resulting solutions are functions of the interest rate and of the ratios m and e . Using these relationships, the interest rate can then be determined as the solution to the following interest parity condition:

$$(1 + r)(1 - F(R^*(r, m, e))) + \mathcal{L}(R^*(r, m, e), R_{EC}^*(r, m, e), r, m, e) = 1 + \bar{r}, \quad (19)$$

where R_{EC}^* solves equation (18) with equality and, for the sake of notational simplicity, the prior return distribution is denoted by $F(R) \equiv \Phi(\sqrt{\alpha}(R - \bar{R}))$ and the expected liquidation value of the bank (per unit of deposits) is expressed in compact notation by $\mathcal{L}(R^*, R_{EC}^*, r, m, e)$.²⁶

However, it is not obvious whether a solution r^* to equation (19) exists and, if it exists, whether it is unique. This is due to the positive dependency of the default probability on the interest rate. As creditors require a higher interest rate given an increase in default risk, the default risk increases further and induces an additional increase in the interest rate, and so on. This mechanism may, under certain conditions, induce a multiplicity of solutions. A thorough treatment of this possibility is beyond the scope of the present paper. I will focus here on the particular case where a unique solution exists in order to examine how the conclusions from the baseline model concerning the liquidity and solvency effects are altered. Intuitively, a sufficient condition for a unique solution is that in equilibrium

²⁴Due to the nature of the sequential service constraint, the probability is given by the ratio of total resources to total claims.

²⁵See Goldstein and Pauzner (2005) for a similar assumption on the payoff structure in their global game version of the Diamond-Dybvig model.

²⁶A more detailed derivation of the condition and the explicit expression for $\mathcal{L}(\cdot)$ is provided in the appendix.

the creditor's marginal loss of an increase in the interest rate due to the higher default risk is smaller than the marginal benefit of such an increase stemming from the higher interest payments and the induced marginal change in the bank's liquidation value. That is, denoting the ratio of marginal loss to marginal benefit by ξ , equation (19) admits a unique solution if²⁷

$$\xi(R^*, R_{EC}^*, r^*) \equiv \xi^* \leq 1. \quad (20)$$

Thus, if conditions (6) and (20) are satisfied, the model has a unique solution which can be written in terms of liquidity and capital ratios as $R^* = R^*(m, e)$, $r^* = r^*(m, e)$, and $R_{EC}^* = R_{EC}^*(m, e)$.

Compared with the baseline model, changes in e and m are now associated with two additional effects due to the endogeneity of the interest rate. Firstly, the interdependency between default risk and interest rate induces a multiplier which magnifies the effects of liquidity and capital ratios. Secondly, changes in e and m affect the liquidation value of the bank and thereby the interest rate and the default risk. I refer to the marginal change in the liquidation value due to marginal changes in e or m as their respective *liquidation value effects*. Intuitively, the liquidation value effect of e has a clear-cut sign. A marginal increase in the capital ratio raises the liquidation value per unit of deposits. This allows to reduce the interest rate and lowers the default risk. But the liquidation value effect of m is ambiguous. Whether a marginal increase in m raises or lowers the liquidation value depends on whether the marginal return from holding more cash exceeds the foregone returns from holding less liquid assets. Whenever the profitable asset's expected return (conditional on the bank defaulting) is sufficiently large, an increase in the liquidity ratio reduces the liquidation value. This leads to a higher interest rate and raises the default risk. And conversely, whenever the expected return is sufficiently low, the liquidation value effect tends to lower both the interest rate and the default risk.

The following proposition summarizes the comparative statics of R^* and r^* .

²⁷Again, the derivation of ξ is provided in the appendix.

PROPOSITION 2.

- (i) R^* and r^* are strictly decreasing in the capital ratio e .
(ii) The derivatives of R^* and r^* with respect to the liquidity ratio m are given by

$$\begin{aligned}\frac{dR^*}{dm} &= \frac{-\frac{\partial\psi/\partial m}{\partial\psi/\partial R^*} - \rho \left(-\frac{\partial\psi/\partial r}{\partial\psi/\partial R^*} \right)}{1 - \xi^*} \quad \text{and} \\ \frac{dr^*}{dm} &= \frac{\frac{\partial\psi/\partial m}{(\partial\psi/\partial r)} \xi^* - \rho}{1 - \xi^*},\end{aligned}\tag{21}$$

which can be either increasing or decreasing, and where the liquidation value effect is reflected by the variable ρ .

Consider the derivatives in (21). The multiplier effect due to the feedback between default risk and interest rate can be seen from the term $1 - \xi^* \in (0, 1)$ in the denominator. Moreover, besides the partial effect on R^* (the first term in the numerator, which can be either positive or negative depending on whether the solvency or the liquidity effect dominates), the liquidation value effect, which is proportional to ρ , enters with a negative sign.²⁸ For example, when $\rho > 0$ (positive liquidation value effect), a marginal increase in m raises the liquidation value. As can be seen from (21), this tends to lower the deposit rate r^* and tends to decrease the default risk. Whenever $\rho < 0$, the effects are reversed. However, even though the sign of the liquidation value effect can be either positive or negative, there exists a close connection between liquidity, solvency, and liquidation value effect.

PROPOSITION 3. *If the liquidity effect dominates the solvency effect, $R^* < (1 + \lambda)(1 - e)$, then the liquidation value effect is positive, $\rho > 0$. Yet, if the liquidation value effect is negative, $\rho < 0$, then the solvency effect dominates the liquidity effect, $R^* > (1 + \lambda)(1 - e)$.*

²⁸ ρ is proportional to the derivative of the liquidation value with respect to m . It is thus a function of endogenous variables and exogenous parameters. For notational simplicity I omit this dependency. The proof of proposition 2 contains the formal definition of ρ .

Propositions 2 and 3 together imply that the dominance of the liquidity effect is a sufficient condition for an increase in the liquidity ratio to effectively reduce the bank's default risk. Note that condition (12) is still the valid condition for assessing the dominance of the liquidity or solvency effect. Conditional on default, the expected return on the bank's asset is low when the liquidity effect dominates. A marginal increase in the liquidity ratio thus increases the liquidation value of the bank because in expected terms the additional cash holdings are worth more than their opportunity cost in terms of foregone investments. As indicated in proposition 3, the converse is not necessarily true. So even if the solvency effect dominates the liquidity effect, the liquidation value effect may still be negative and counteract the effect of a higher liquidity ratio on the likelihood of bank default.

With respect to the dominance of either solvency or liquidity effect, recall that in the baseline model, condition (12) was rewritten in terms of the interest rate/capital ratio to obtain a condition purely in terms of the exogenous parameters. Since the interest rate is now determined endogenously, one may therefore rewrite the condition in terms of the exogenous risk-free rate $\bar{r}$. The intuition remains broadly similar compared with the baseline model. If the risk-free rate is rather large, the interest rate on bank deposits must be large as well. This implies that the solvency point R_s increases strongly in m and pushes up the default risk. Thus, as shown in the proof of proposition 2, there is a critical value $\hat{r}^m(e)$ such that if $\bar{r} > \hat{r}^m(e)$, then $R^* > (1 - e)(1 + \lambda)$ and the solvency effect dominates the liquidity effect. Conversely, whenever $\bar{r} < \hat{r}^m(e)$, the liquidity effect is dominant and a higher liquidity ratio can be used to effectively reduce the bank's default risk. In practice, since the risk-free rate correlates positively with the interest rate set by the monetary authority, central bank actions may influence the effectiveness of bank regulatory measures. In particular, in low interest environments, a higher liquidity ratio may be effective in reducing bank default risk for a given capital ratio, whereas in high interest environments, additional increases in the capital ratio should accompany increases in the liquidity ratio to increase $\hat{r}^m(e)$ so that the default risk can be stabilized.

However, the relationship between liquidity and capital regulation that was visible from figure 1 becomes more complicated. By

inverting $\hat{r}^m(e)$, one can still derive a critical capital ratio such that the liquidity effect dominates once the capital ratio exceeds this critical value. But it is not obvious how the liquidity ratio should be adjusted below this critical value. The existence of the liquidation value effect may in fact undermine the relationship between capital and liquidity regulation. There may be regions in e - m space where the liquidation value effect outweighs the solvency effect and conversely.

Hence, in the extreme case of no risk premium, liquidity regulation just has to take into account whether the capital ratio e is above or below the critical level e^m . But in the opposite extreme of a perfectly risk-reflecting interest rate, regulators have to take into account either the intricate relationship between liquidation value and solvency effects or rely only on the sufficient condition and impose a sufficiently high capital ratio to be sure to get around the double-edged-sword problem of liquidity regulation.

5.2 *Endogenous Fire-Sale Rate*

Next, return to the baseline model with exogenous interest rate r , but consider the effect of an endogenous fire-sale rate λ . With respect to the origin of the fire-sale rate, Rochet and Vives (2004, p. 1122) stress, in particular, the presence of asymmetric information. They point out that in a market plagued by informational asymmetries, a bank may have essentially two reasons for selling assets—an urgent liquidity need or the desire to offload bad assets. Having information inferior to that of the bank, market participants are unable to distinguish between the two reasons and offer to purchase assets only against a discount. Modeling this informational friction in greater detail would allow to express the fire-sale discount in terms of (exogenously given) parameters describing, for example, the fraction of good and bad assets in the economy or informational assumptions. While this would allow for more complex and detailed comparative statics, it would not necessarily render the fire-sale discount endogenously dependent on the failure point R^* . However, such a dependency generically occurs if the fire-sale rate is influenced by aggregate sales. This could, for example, reflect a situation where a bank's holdings of a particular asset class is large relative to the totally traded volume in the market. In line with the discussion in

Morris and Shin (2004), it may also proxy the situation of a certain group of financial institutions with similar asset positions operating in a market with a downward-sloping residual demand curve. As a consequence, the larger the amount withdrawn, the more assets have to be sold and the more deflated the asset price will be. A straightforward way to capture this feedback is to assume that the fire-sale discount rate depends negatively on the volume of sales,

$$\lambda = \zeta(\max\{x - m, 0\}).$$

I assume further that the price equals the fundamental price if no fire sales occur, $\zeta(0) = 0$. To ensure that the model behaves well, I additionally assume that prices, even though they may be quite small in the presence of maximal fire sales, are never zero, $\zeta(1) = \bar{\lambda} < \infty$, and that the rate of price change is bounded $\zeta' \in (0, \kappa]$. Finally, for computational simplicity, $\zeta'' = 0$.

Since a larger amount of fire sales lowers the asset's price by pushing up λ , the strategic complementarities in the withdrawal decisions of fund managers are strengthened. This requires imposing a more stringent condition than condition (6) to guarantee the uniqueness of the monotone global game equilibrium. The respective sufficient condition is derived in the appendix and is given by

$$\beta \geq \frac{\alpha^2}{2\pi} \left(\frac{\bar{\lambda}D + \kappa(D - M)}{I} \right)^2. \quad (22)$$

Given that condition (22) holds, the threshold value R^* can be used to compute the probability of a bank default. Moreover, as the following proposition makes clear, under certain conditions, the qualitative results from the baseline model may still hold and there may exist a critical interest rate/capital ratio where the relation between liquidity and solvency effect changes.

PROPOSITION 4. *If the fire-sale rate changes not too strongly, $\kappa \leq \frac{\sqrt{2\beta\pi}}{2\alpha}$, there exists a threshold function $\ell(m, e)$ with $\partial\ell/\partial m < 0$ and $\partial\ell/\partial e < 0$ such that the liquidity effect dominates the solvency effect if and only if*

$$R^* < \ell(m, e). \quad (23)$$

As in section 3, condition (23) can be used to derive a critical value $\hat{r}(e)$ such that the liquidity effect dominates the solvency effect if $r \leq \hat{r}(e)$. Liquidity and solvency effects are thus preserved also in the presence of an adverse feedback from withdrawals to prices.²⁹

However, once the assumption in proposition 4 fails to hold and the fire-sale rate responds too strongly to the underlying withdrawal pressure, the monotone relationship between the failure point and the dominance of either solvency or liquidity effect may break down; a non-monotone relationship may emerge with more than one critical value demarcating regions in e - m space where the liquidity or the solvency effect dominates.

Given that the assumption in the foregoing proposition is met, the question arises to which extent regulatory measures exert influence over the expected fire-sale price? This would indicate the existence of an important effect of regulatory measures. Not only may regulation work to directly increase the liquidity of certain financial institutions, it may also improve funding terms *ex ante*, by affecting expected asset prices and thereby liquidity and risk premia that would have to be incurred when assets are used, for example, as collateral in other transactions.

Using equation (4), the expected fire-sale price can be written as

$$\begin{aligned} & \mathbf{E} \left(\frac{R}{1 + \zeta(\max\{x - m, 0\})} \right) \\ &= \int_{-\infty}^{\frac{1}{\sqrt{\beta}} z(\Phi^{-1}(m)/\sqrt{\beta}; R^*)} R (1 + \zeta(\Phi(z(R; R^*)) - m))^{-1} dF(R) \\ &+ \int_{\frac{1}{\sqrt{\beta}} z(\Phi^{-1}(m)/\sqrt{\beta}; R^*)}^{\infty} R dF(R), \end{aligned}$$

where $z(R; R^*) \equiv \frac{\alpha}{\sqrt{\beta}}(R^* - \bar{R}) + \sqrt{\beta}(R^* - R) - \sqrt{\frac{\alpha+\beta}{\beta}}\Phi^{-1}(\gamma)$.

²⁹Note that even though $\ell(m, e)$ depends on both liquidity and capital ratio, the critical value $\hat{r}$ does not depend on m anymore since it is, by definition, the value of r where R^* ceases to depend on m .

By differentiating this expression with respect to e , it is straightforward to show that the expected fire-sale price strictly increases in the capital ratio. This implies that higher capital requirements lower *ex ante* expected losses in case fire sales become indeed necessary. The effects of the liquidity ratio, however, are ambiguous. Firstly, a higher liquidity ratio raises the fraction of withdrawals needed to trigger fire sales. Secondly, a higher liquidity ratio also affects the fraction of withdrawals directly, as it causes a change in the failure point. These two effects work in the same direction once the liquidity effect outweighs the solvency effect. In such a situation, a higher liquidity ratio depresses fire-sale discounts and helps to mitigate losses that may eventually drive the bank into default. But once the solvency effect dominates, the effects work in opposite directions. While higher cash holdings may help the bank to delay fire sales, they drive up the solvency risk which, in turn, raises fund managers' incentives to run to the bank. Thereby, the impact on the expected fire-sale discount becomes weaker and may eventually turn around, thereby driving up the expected fire-sale rate.

6. Conclusion

The paper revisits the well-known banking model by Rochet and Vives and analyzes complementarity and substitutability between liquidity and capital regulation. Importantly, the analysis fully takes into account the adjustments in the balance sheet needed to accommodate the specific variations in the regulatory ratios. It demonstrates that liquidity requirements, like the Basel Accord's newly proposed LCR, are only effective in strengthening the resilience of a bank once the liquidity effect outweighs the solvency effect. Albeit rather stylized, the model points to some important aspects in the design of regulatory rules. Firstly, besides strengthening the bank's ability to withstand sudden funding roll-offs, liquidity requirements may have a direct impact on the solvency of the bank. Secondly, whether liquidity and capital requirements are substitutes or complements depends on the level of the capital ratio. For higher liquidity ratios to effectively lower bank default risk, a sufficiently strong capital buffer above a certain critical level is required. Finally, the double-edged-sword nature of liquidity regulation obtains also in more general environments where interest rates carry risk premia

or fire sales feed back into prices. Under such circumstances, however, the relationship between capital and liquidity regulation may become more complex.

The paper leaves several questions open that remain to be addressed in future research. Theoretically, since the Rochet-Vives model is a partial equilibrium model, a deeper understanding of the interactions between liquidity and capital regulation may be obtained by embedding the coordination problem of fund managers and creditors into a (general equilibrium) framework where the bank's composition of asset and liability sides are endogenous.³⁰ Moreover, the analysis in the present paper raises the empirical question whether the liquidity or the solvency effect dominates in practice and how large the critical capital ratio e^m is such that increases in liquidity ratios can reduce bank default risk.

Appendix

Proof of Proposition 1

If $R^* = R_s$, the failure point is given explicitly by equation (9) and it is straightforward to compute

$$\begin{aligned} \frac{\partial R_s}{\partial e} &= -\frac{(1-m)(1+r)}{1-m(1+r)} < 0 \quad \text{and} \quad \frac{\partial R_s}{\partial m} = \frac{(1+r)(1-e)r}{(1-m(1+r))^2} \geq 0 \\ \text{and} \quad \frac{\partial R_s}{\partial r} &= \frac{(1-e)(1-m)}{(1-m(1+r))^2} \geq 0. \end{aligned}$$

If $R^* > R_s$, the failure point is given by the solution to the implicit function (11). Since $\beta \geq \beta_0$ and the equilibrium is unique,

³⁰Recent papers by Liu and Mello (2011) and Daniëls and König (2013) study financial institutions' choice of the asset side, whereas other work by Eisenbach (2013), Szkup (2013), or König and Pothier (2014) consider the determination of the liability side in global game bank-run models. What is missing so far from the literature is a treatment of an endogenous determination of both the asset and liability side in a coordination game setup.

$$\begin{aligned} \frac{\partial \psi(R^*, m, e, r)}{\partial R} &= (1 - m(1 + r)) - \frac{\alpha\lambda(1 - e)(1 + r)}{\sqrt{\beta}}\phi \\ &\times \left(\frac{\alpha}{\sqrt{\beta}}(R^* - \bar{R}) - \sqrt{\frac{\alpha + \beta}{\beta}}\Phi^{-1}(\gamma) \right) > 0. \end{aligned}$$

For any $z \in \{m, e, r\}$, by the implicit function theorem,

$$\frac{\partial R^*}{\partial z} < 0 \Leftrightarrow -\frac{\frac{\partial \psi}{\partial z}}{\frac{\partial \psi}{\partial R}} < 0 \Leftrightarrow \frac{\partial \psi}{\partial z} > 0. \quad (24)$$

The partial derivative with respect to the e is given by

$$\begin{aligned} \frac{\partial \psi(R^*, m, e, r)}{\partial e} &= (1 + r)(1 - m) + \lambda(1 + r) \\ &\times \left(\Phi \left(\frac{\alpha}{\sqrt{\beta}}(R^* - \bar{R}) - \sqrt{\frac{\alpha + \beta}{\beta}}\Phi^{-1}(\gamma) \right) - m \right) > 0 \end{aligned}$$

and by equation (24), $\frac{\partial R^*}{\partial e} < 0$. The partial derivative with respect to r is given by

$$\begin{aligned} \frac{\partial \psi}{\partial r} &= -mR^* - (1 - e)(1 - m) - \lambda(1 - e) \\ &\times \left(\Phi \left(\frac{\alpha}{\sqrt{\beta}}((1 - e)(1 + \lambda) - \bar{R}) - \sqrt{\frac{\alpha + \beta}{\beta}}\Phi^{-1}(\gamma) \right) - m \right) < 0, \end{aligned}$$

and it follows from equation (24) that $\frac{\partial R^*}{\partial r} > 0$. The partial derivative with respect to the m is given by

$$\frac{\partial \psi(R^*, m, e, r)}{\partial m} \leq 0 \Leftrightarrow R^* \geq (1 - e)(1 + \lambda).$$

Using (24), it follows that $\frac{\partial R^*}{\partial m} < 0$ if and only if $R^* < (1 - e)(1 + \lambda)$.

Derivation of Iso-Risk Curves

The derivation of iso-risk curves proceeds in two steps. First, the critical level $e^m(r)$ and the corresponding risk level R_m^* are derived. Second, the properties of iso-risk curves for risk levels $R_k^* \neq R_m^*$ are shown.

- (i) The critical level $e^m(r)$ and the associated risk level R_m^* constitute the simultaneous solution to the two equations

$$\psi(R_m^*, m, e^m, r) = 0 \quad \text{and} \quad R_m^* - (1 - e^m)(1 + \lambda) = 0.$$

The explicit solutions are given by

$$R_m^* = \bar{R} + \frac{\sqrt{\beta}}{\alpha} \Phi^{-1} \left(\frac{\lambda - r}{\lambda(1 + r)} \right) + \sqrt{\frac{\alpha + \beta}{\alpha^2}} \Phi^{-1}(\gamma)$$

and

$$e^m = 1 - \frac{\bar{R} + \frac{\sqrt{\beta}}{\alpha} \Phi^{-1} \left(\frac{\lambda - r}{\lambda(1 + r)} \right) + \sqrt{\frac{\alpha + \beta}{\alpha^2}} \Phi^{-1}(\gamma)}{1 + \lambda}.$$

Since e^m is independent of m , it constitutes a horizontal line in e - m space. Along this line, variations in the liquidity ratio m do not exert any effect on the risk level R_m^* .

- (ii) For fixed risk levels $R_k^* \neq R_m^*$, the iso-risk curves in e - m space are implicitly defined by $\psi(R_k^*, m, e, r) = 0$. Since $\partial\psi/\partial e > 0$, a capital ratio $e > e^m$ implies a risk level $R_k^* < R_m^*$ and vice versa for $e < e^m$. The slope of the iso-risk curve is given by

$$\left. \frac{de}{dm} \right|_{R^*=R_k^*} = - \frac{\partial\psi/\partial m}{\partial\psi/\partial e},$$

which is positive for any $R_k^* > R_m^*$ ($e < e^m$) and negative for any $R_k^* < R_m^*$ ($e > e^m$). Moreover,

$$\left. \frac{d^2e}{dm^2} \right|_{R^*=R_k^*} = - \frac{2\partial^2\psi/\partial m\partial e}{\partial\psi/\partial e} \frac{de}{dm}.$$

Since $\frac{\partial^2\psi}{\partial e\partial m} < 0$, the iso-risk curves are concave for $R_k^* < R_m^*$ (i.e., $e > e^m$) and convex for $R_k^* > R_m^*$ (i.e., $e < e^m$).

Derivation of Interest Parity Condition (19)

With probability $1 - F(R^*)$, creditors obtain D independent of whether fund managers withdraw or not.

If the bank fails early, $R < R_{EC}$, its resources are insufficient to meet withdrawals. Creditors face a sequential service constraint, and the likelihood of receiving D is given by $(RI(1 + \lambda)^{-1} + M)/xD$. By the law of large numbers, the probability of withdrawing—i.e., of getting a signal below t^* conditional on R —is given by

$$\Phi(\sqrt{\beta}(t^*(R^*) - R)) \equiv x(t^*(R^*), R) \equiv x(R^*, R).$$

Creditors whose funds are rolled over go away empty handed. Hence, if $R < R_{EC}$, a creditor expects to receive

$$\begin{aligned} & \int_{\underline{R}}^{R_{EC}} \frac{(RI(1 + \lambda)^{-1} + M)}{x(R^*, R)D} x(R^*, R) D \, dF(R) \\ &= (1 + r)D_0 \int_{\underline{R}}^{R_{EC}} \left[\frac{R(1 - m(1 + r))}{(1 + \lambda)(1 - e)(1 + r)} + m \right] dF(R), \end{aligned}$$

where $\underline{R} \equiv -\frac{(1+\lambda)m(1+r)(1-e)}{1-(1+r)m}$ is the value of R such that the probability of a creditor getting served is at least zero.

Next, consider the case $R \in (R_{EC}, R^*)$. Conditional on R , a creditor withdraws with probability $x(R^*, R)$ and obtains D (available liquidity in this range is sufficient to meet withdrawals). The bank fails at date 2, so that creditors who roll over (with probability $(1 - x(R^*, R))$) become residual claimants. The residual wealth of the bank (taking the losses due to early fire sales into account) is given by

$$RI + (1 + \lambda)M - (1 + \lambda)x(R^*, R)D.$$

This is divided on a pro rata basis among the remaining $(1 - x)D$ liabilities. So a creditor obtains a fraction

$$\frac{(RI + (1 + \lambda)M - (1 + \lambda)x(R^*, R)D)}{(1 - x(R^*, R))D}$$

of his original claim D .

That is, in case $R \in (R_{EC}, R^*)$, he expects to receive

$$D \int_{R_{EC}}^{R^*} x(R^*, R) dF(R) \\ + \int_{R_{EC}}^{R^*} \left[\frac{(RI + (1 + \lambda)M - (1 + \lambda)x(R^*, R)D)}{(1 - x(R^*, R))D} (1 - x(R^*, R))D \right] \\ \times dF(R),$$

which can be written as

$$(1 + r)D_0 \int_{R_{EC}}^{R^*} x(R^*, R) dF(R) \\ + (1 + r)D_0 \int_{R_{EC}}^{R^*} \left[\frac{R(1 - m(1 + r))}{(1 - e)(1 + r)} + (1 + \lambda)(m - x(R^*, R)) \right] \\ \times dF(R).$$

Putting everything together, the interest parity condition can be written as

$$(1 - F(R^*)) (1 + r)D_0 \\ + (1 + r)D_0 \int_{\underline{R}}^{R_{EC}} \left[\frac{R(1 - m(1 + r))}{(1 + \lambda)(1 - e)(1 + r)} + m \right] dF(R) \\ + (1 + r)D_0 \int_{R_{EC}}^{R^*} x(R^*, R) dF(R) \\ + (1 + r)D_0 \int_{R_{EC}}^{R^*} \left[\frac{R(1 - m(1 + r))}{(1 - e)(1 + r)} + (1 + \lambda)(m - x(R^*, R)) \right] \\ \times dF(R) = (1 + \bar{r})D_0, \quad (25)$$

which can be expressed in compact form as in equation (19) by dividing both sides by D_0 and using the abbreviation

$$\begin{aligned}
 & \mathcal{L}(R^*, R_{EC}, r, m, e) \\
 & \equiv (1+r) \left[\int_{\underline{R}}^{R_{EC}} \left[\frac{R(1-m(1+r))}{(1+\lambda)(1-e)(1+r)} + m \right] dF(R) \right] \\
 & + (1+r) \left[\int_{R_{EC}}^{R^*} x(R^*, R) dF(R) \right. \\
 & \left. + \int_{R_{EC}}^{R^*} \left[\frac{R(1-m(1+r))}{(1-e)(1+r)} + (1+\lambda)(m-x(R^*, R)) \right] dF(R) \right].
 \end{aligned}$$

Derivation of Uniqueness Condition (20)

R^* and R_{EC}^* are determined as functions of r , m , and e from

$$\psi(R^*, r, m, e) = 0 \quad (26)$$

$$\frac{R_{EC}^*(1-m(1+r))}{(1+\lambda)(1-e)} + (1+r)(m-x(R^*, R_{EC}^*)) = 0. \quad (27)$$

A sufficient condition for a stable/unique equilibrium is that the left-hand side of the interest parity condition strictly increases in r when evaluated at the equilibrium values r^* , R^* , and R_{EC}^* . Taking the derivative of equation (25) with respect to r and evaluating it at the equilibrium values, this condition becomes

$$\begin{aligned}
 & (1-F(R^*)) + m \int_{\underline{R}}^{R_{EC}^*} \left[-\frac{R}{(1+\lambda)(1-e)} + 1 \right] dF(R) \\
 & + (1+\lambda)m \int_{R_{EC}^*}^{R^*} \left[-\frac{R}{(1+\lambda)(1-e)} + 1 \right] dF(R) \\
 & - \lambda \int_{R_{EC}^*}^{R^*} x(R^*, R) dF(R) \\
 & - (1+\lambda)(1+r^*) \int_{R_{EC}^*}^{R^*} x_{R^*}(R^*, R) \frac{\partial R^*}{\partial r} \Big|_{r=r^*} dF(R) > 0.
 \end{aligned}$$

The latter can be rewritten as

$$\begin{aligned} & \xi(R^*, R_{EC}^*, r^*) \\ & \equiv \frac{\overbrace{(1 + \lambda)(1 + r^*) \int_{R_{EC}^*}^{R^*} x_{R^*}(R^*, R) \left. \frac{\partial R^*}{\partial r} \right|_{r=r^*} dF(R)}^{\text{marginal loss due to higher default risk}}}{\underbrace{1 - F(R^*) - m \left[\int_{\underline{R}}^{R_{EC}^*} \left[\frac{R}{(1 + \lambda)(1 - e)} - 1 \right] dF(R) + \lambda \int_{R_{EC}^*}^{R^*} \left[\frac{R}{(1 + \lambda)(1 - e)} - 1 + \frac{x(R^*, R)}{m} \right] dF(R) \right]}_{\text{marginal benefit due to higher interest and changed liquidation value}}} < 1. \end{aligned}$$

Proof of Proposition 2

I use Cramer's rule to compute the derivatives of R^* and r^* with respect to e and m . Totally differentiating the system of equations consisting of (25), (27), and (28) and casting the results into matrix form gives

$$\mathbf{J}\mathbf{v} + \mathbf{B}\mathbf{u} = 0 \quad \text{where} \quad \mathbf{v}^T = (dr^* \quad dR^* \quad dR_{EC}^*), \quad \mathbf{u}^T = (dm \quad de).$$

The elements of the 3×3 Jacobian matrix, evaluated at the equilibrium, are given by

$$\begin{aligned} j_{11} &= (1 - F(R^*)) + m \int_{\underline{R}}^{R_{EC}^*} \left[-\frac{R}{(1 + \lambda)(1 - e)} + 1 \right] dF(R) \\ &\quad + (1 + \lambda)m \int_{R_{EC}^*}^{R^*} \left[-\frac{R}{(1 + \lambda)(1 - e)} + 1 \right] dF(R) \\ &\quad - \lambda \int_{R_{EC}^*}^{R^*} x(R^*, R) dF(R) \\ &= (1 - F(R^*)) + \frac{m(F(R_{EC}^*) - F(\underline{R}))}{(1 + \lambda)(1 - e)} \left((1 + \lambda)(1 - e) \right. \\ &\quad \left. - \mathbb{E} \left(\tilde{R} \mid R \in (\underline{R}, R_{EC}^*) \right) \right) \end{aligned}$$

$$\begin{aligned}
 & + \frac{m(F(R^* - F(R_{EC}^*)))}{(1 + \lambda)(1 - e)} \left((1 + \lambda)(1 - e) \right. \\
 & \left. - \mathbb{E} \left(\tilde{R} \middle| R \in (R_{EC}, R^*) \right) \right) - \lambda \int_{R_{EC}^*}^{R^*} x(R^*, R) dF(R), \\
 j_{12} & = -(1 + \lambda)(1 + r^*) \int_{R_{EC}}^{R^*} x_{R^*}(R^*, R) dF(R) < 0, \quad j_{13} = 0, \\
 j_{21} & = \frac{\partial \psi}{\partial r} \bigg|_{R^*, r=r^*} < 0, \quad j_{22} = \frac{\partial \psi}{\partial R^*} \bigg|_{R^*, r=r^*} > 0, \quad j_{23} = 0, \\
 j_{31} & = \frac{-mR_{EC}^*}{(1 + \lambda)(1 - e)} + m - x(R^*, R_{EC}^*) < 0, \\
 j_{32} & = -(1 + r^*)x_{R^*}(R^*, R_{EC}^*) < 0, \\
 j_{33} & = \frac{1 - m(1 + r^*)}{(1 + \lambda)(1 - e)} - (1 + r^*)x_{R_{EC}^*}(R^*, R_{EC}^*) > 0.
 \end{aligned}$$

Given the condition for a unique equilibrium and since $j_{12} < 0$, it follows that $j_{11} > 0$. The determinant of the Jacobian is given by $|\mathbf{J}| = j_{33}(j_{11}j_{22} - j_{21}j_{12}) > 0$, which can be written as $|\mathbf{J}| = j_{11}j_{22}j_{33}(1 - \xi^*)$.

The elements of the 3×2 matrix B are given by

$$\begin{aligned}
 b_{11} & = \frac{(1 + r^*)(F(R_{EC}^* - F(\underline{R})))}{(1 + \lambda)(1 - e)} \left((1 + \lambda)(1 - e) \right. \\
 & \left. - \mathbb{E} \left(\tilde{R} \middle| R \in (\underline{R}, R_{EC}^*) \right) \right) \\
 & + \frac{(1 + r^*)(F(R^* - F(R_{EC}^*)))}{(1 - e)} \left((1 + \lambda)(1 - e) \right. \\
 & \left. - \mathbb{E} \left(\tilde{R} \middle| R \in (R_{EC}^*, R^*) \right) \right) \geq 0, \\
 b_{21} & = \frac{\partial \psi}{\partial m} \bigg|_{R^*, r=r^*} \geq 0, \quad b_{31} = -(1 + r^*) \left(\frac{R_{EC}^*}{(1 + \lambda)(1 - e)} - 1 \right) \geq 0, \\
 b_{12} & = \int_{\underline{R}}^{R_{EC}^*} \frac{R(1 - m(1 + r^*))}{(1 + \lambda)(1 - e)^2} dF(R) \\
 & + \int_{R_{EC}^*}^{R^*} \frac{R(1 - m(1 + r^*))}{(1 - e)^2} dF(R) > 0
 \end{aligned}$$

$$b_{22} = \frac{\partial \psi}{\partial e} \bigg|_{R^*, r=r^*} > 0, \quad b_{23} = \frac{(1+r^*)R_{EC}^*}{(1+\lambda)(1-e)^2} > 0.$$

Using Cramer's rule, it is straightforward to establish part (i) of the proposition.

$$\frac{dR^*}{de} = \frac{\begin{vmatrix} j_{11} & -b_{12} & 0 \\ j_{21} & -b_{22} & 0 \\ j_{31} & -b_{32} & j_{33} \end{vmatrix}}{|\mathbf{J}|} < 0 \quad \text{and} \quad \frac{dr^*}{de} = \frac{\begin{vmatrix} -b_{12} & j_{12} & 0 \\ -b_{22} & j_{22} & 0 \\ -b_{32} & j_{32} & j_{33} \end{vmatrix}}{|\mathbf{J}|} < 0$$

With respect to part (ii), define

$$\rho \equiv \frac{b_{11}}{j_{11}},$$

whose sign is determined by b_{11} as $j_{11} > 0$. b_{11} is the liquidation value effect of m since it is the marginal change in the liquidation value with respect to m . One obtains

$$\frac{dR^*}{dm} = \frac{\begin{vmatrix} j_{11} & -b_{11} & 0 \\ j_{21} & -b_{21} & 0 \\ j_{31} & -b_{31} & j_{33} \end{vmatrix}}{|\mathbf{J}|} \quad \text{and} \quad \frac{dr^*}{dm} = \frac{\begin{vmatrix} -b_{11} & j_{12} & 0 \\ -b_{21} & j_{22} & 0 \\ -b_{31} & j_{32} & j_{33} \end{vmatrix}}{|\mathbf{J}|},$$

which can be rewritten as

$$\frac{dR^*}{dm} = \frac{\frac{-b_{21}}{j_{22}} + \frac{j_{21}b_{11}}{j_{11}j_{22}}}{1 - \xi^*} \quad \text{and} \quad \frac{dr^*}{dm} = \frac{\frac{-b_{11}}{j_{11}} + \frac{j_{21}b_{21}j_{12}}{j_{21}j_{11}j_{22}}}{1 - \xi^*},$$

which, by using the respective expressions given above, yields the expressions in (21).

Recall from condition (12) and equation (13) that $R^* < (1 + \lambda)(1 - e)$ if and only if $r^* < r^m(e)$. Evaluating equation (25) at $r^m(e)$ and solving for $\bar{r}$ yields the value of the risk-free $\hat{r}^m$ that induces $r^* = r^m(e)$. As the interest rate and the default risk strictly increase in $\bar{r}$, if $\bar{r} < \hat{r}^m$, we have $r^* < r^m(e)$ and the liquidity effect dominates the solvency effect.

Proof of Proposition 3

Define $a_0 \equiv \frac{(1+r^*)(F(R_{EC})-F(\underline{R}))}{(1+\lambda)(1-e)j_{11}} > 0$ and $a_1 \equiv \frac{(1+r^*)(F(R^*)-F(R_{EC}))}{(1-e)j_{11}} > 0$.

Suppose first that the liquidity effect dominates the solvency effect. From proposition 1, this occurs if and only if $R^* < (1-e)(1+\lambda)$. This implies

$$\begin{aligned} & 0 > (a_0 + a_1)(R^* - (1-e)(1+\lambda)) \\ & > a_0 ((F(R_{EC}^*) - F(\underline{R}))R^* - (1-e)(1+\lambda)) \\ & \quad + a_1 ((F(R^*) - F(R_{EC}^*))R^* - (1-e)(1+\lambda)) \\ & > a_0 \left(\int_{\underline{R}}^{R_{EC}^*} R dF(R) - (1-e)(1+\lambda) \right) \\ & \quad + a_1 \left(\int_{R_{EC}^*}^{R^*} R dF(R) - (1-e)(1+\lambda) \right) \\ & = -\rho. \end{aligned}$$

This establishes that if the liquidity effect dominates the solvency effect, then the liquidation value effect is positive. It follows immediately that the contrapositive of this claim—i.e., if the liquidation value effect is negative, then the solvency effect dominates the liquidity effect—is also true.

Derivation of Uniqueness Condition (22)

Since the dependency of the fire-sale price on aggregate withdrawals strengthens the strategic complementarities between the fund managers, the required uniqueness condition becomes more stringent. Inserting $\zeta(\cdot)$ into equation (11) and differentiating with respect to R^* gives

$$\begin{aligned} & \frac{d\psi(R^*, m, e, r, \zeta(R^*, m))}{dR^*} \\ & = (1 - m(1+r)) - \frac{\lambda\alpha}{\sqrt{\beta}}\phi\left(\frac{\alpha}{\sqrt{\beta}}(R^* - \bar{R}) - \sqrt{\frac{\alpha+\beta}{\beta}}\Phi^{-1}(\gamma)\right) \\ & \quad \times (1-e)(1+r) \end{aligned}$$

$$\begin{aligned}
& -\zeta'(1-e)(1+r)\phi\left(\frac{\alpha}{\sqrt{\beta}}(R^* - \bar{R}) - \sqrt{\frac{\alpha+\beta}{\beta}}\Phi^{-1}(\gamma)\right)\frac{\alpha}{\sqrt{\beta}} \\
& \times \left(\Phi\left(\frac{\alpha}{\sqrt{\beta}}(R^* - \bar{R}) - \sqrt{\frac{\alpha+\beta}{\beta}}\Phi^{-1}(\gamma)\right) - m\right).
\end{aligned}$$

Since $\zeta' > 0$, the sufficient condition for the uniqueness of the threshold equilibrium has to be modified by using the assumptions $\lambda < \bar{\lambda}$ and $\zeta' < \kappa$ to yield

$$\sqrt{\beta} > \frac{\alpha}{\sqrt{2\pi}} \frac{(1-e)(1+r)(\bar{\lambda} + \kappa(1-m))}{1-m(1+r)}.$$

Proof of Proposition 4

Inserting $\zeta(\cdot)$ into equation (11) and differentiating with respect to m yields

$$\frac{\partial\psi}{\partial m} \geq 0 \quad \Leftrightarrow \quad R^* \leq \tilde{\ell}(R^*, e, m), \quad (28)$$

where $\tilde{\ell}(R^*, e, m) \equiv (1-e)(1+\lambda + \zeta'(\Phi(\frac{\alpha}{\sqrt{\beta}}(R^* - \bar{R}) - \sqrt{\frac{\alpha+\beta}{\beta}}\Phi^{-1}(\gamma)) - m))$. Equation (29) can be solved for a unique critical value R^* if $\kappa \leq \frac{\sqrt{2\beta\pi}}{2\alpha}$. As long as this condition holds, the left-hand side of equation (29) increases faster than the right-hand side and there exists a unique value $\ell(m, e)$ such that $\partial\psi/\partial m > 0$ (the liquidity effect dominates) if and only if $R^* < \ell(\cdot)$. The derivatives of ℓ with respect to m and e can be easily found by applying the implicit function theorem to $R^* - \tilde{\ell}(R^*, m, e) = 0$. Evaluating equation (11) at $R^* = \ell(m, e)$ yields

$$\psi(\ell(m, e), m, e, r) = \tilde{\psi}(e, r) = 0,$$

which is independent of m since ℓ is exactly equal to the value of R^* where the derivative of $\psi(\cdot)$ with respect to m vanishes. Equation $\tilde{\psi}(e, r) = 0$ can thus be solved for $r = \hat{r}(e)$, and since $\psi(\cdot)$ is strictly decreasing in r , it follows that for any $r < \hat{r}(e)$, $R^* < \ell$ and the liquidity effect dominates the solvency effect.

References

- Basel Committee on Banking Supervision. 2008. "Principles for Sound Liquidity Risk Management and Supervision." (June).
- . 2013. "Basel III: The Liquidity Coverage Ratio and Liquidity Risk Monitoring Tools." (January).
- . 2014. "Basel III: The Net Stable Funding Ratio (Consultative Document)." (January).
- Chevalier, J., and G. Ellison. 1997. "Risk Taking by Mutual Funds as a Response to Incentives." *Journal of Political Economy* 105 (6): 1167–1200.
- . 1999. "Are Some Mutual Fund Managers Better than Others? Cross-Sectional Patterns in Behavior and Performance." *Journal of Finance* 54 (3): 875–99.
- Daniëls, T. R., and P. J. König. 2013. "Illiquidity Risk and Liquidity Regulation." Mimeo.
- Diamond, D. W., and P. H. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (3): 401–19.
- Eisenbach, T. M. 2013. "Rollover Risk as Market Discipline: A Two-Sided Inefficiency." Staff Report No. 597, Federal Reserve Bank of New York.
- Goldstein, I., and A. Pauzner. 2005. "Demand Deposit Contracts and the Probability of Bank Runs." *Journal of Finance* 60 (3): 1293–1328.
- Goodhart, C. A. E. 1999. "Myths about the Lender of Last Resort." *International Finance* 2 (3): 339–60.
- . 2009. "Liquidity and Money Market Operations: A Proposal." In *The Regulatory Response to the Financial Crisis*, ed. C. Goodhart, 59–84. Edward Elgar Publishing.
- Heinemann, F. 2005. "Die Theorie globaler Spiele: Private Information als Mittel zur Vermeidung multipler Gleichgewichte." *Journal für Betriebswirtschaftslehre* 55 (3): 209–41.
- Hellwig, C. 2002. "Public Information, Private Information, and the Multiplicity of Equilibria in Coordination Games." *Journal of Economic Theory* 107 (2): 191–222.
- König, P. J., and D. Pothier. 2014. "Asymmetric Information and Roll-over Risk." Technical Report No. 1364, DIW Berlin.

- Liu, X., and A. S. Mello. 2011. "The Fragile Capital Structure of Hedge Funds and the Limits to Arbitrage." *Journal of Financial Economics* 102 (3): 491–50.
- Morris, S., and H. S. Shin. 2001. "Rethinking Multiple Equilibria in Macroeconomic Modeling." In *NBER Macroeconomics Annual 2000*, Vol. 15, 139–82. National Bureau of Economic Research, Inc., and MIT Press.
- . 2003. "Global Games: Theory and Application." In *Advances in Economics and Econometrics: Theory and Applications*, ed. M. Dewatripont, L. Hansen, and S. Turnovsky, 56–114. Cambridge University Press.
- . 2004. "Liquidity Black Holes." *Review of Finance* 8 (1): 1–18.
- Rochet, J.-C., and X. Vives. 2004. "Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?" *Journal of the European Economic Association* 2 (6): 1116–47.
- Szkup, M. 2013. "Optimal Debt Maturity Structure, Rollover Risk and Strategic Uncertainty." Mimeo (March).

How Much Does the Public Know about the ECB's Monetary Policy? Evidence from a Survey of Dutch Households*

Carin van der Cruijssen,^a David-Jan Jansen,^a and
Jakob de Haan^{a,b,c}

^aDe Nederlandsche Bank

^bUniversity of Groningen

^cCESifo

Does the general public know what central banks do? Is this knowledge relevant? Based on a survey of Dutch households, we find that knowledge about the European Central Bank's objectives is far from perfect. A weak desire to be informed and unawareness among respondents about their limited knowledge are barriers for improving general understanding of monetary policy. However, our results also show that more intensive use of information improves respondents' understanding of the ECB, suggesting that communication may improve general knowledge. Finally, we find that knowledge about monetary policy objectives contributes to an individual's ability to form inflation expectations.

JEL Codes: D12, D84, E52, E58.

*The authors thank Andreas Busch, Maria Demertzis, Michael Ehrmann, Ivan Jaccard, Gabriele Galatti, Joris Knoben, Bartosz Makowiak, Maarten van Rooij, Jiří Slačálek, and Federica Teppa for useful discussions. Three anonymous referees provided many helpful suggestions. We also received insightful comments from participants at the 2010 ASSA meetings and seminars at the European Central Bank, the Deutsche Bundesbank, Göttingen University, the University of Salamanca, the International Monetary Fund, and De Nederlandsche Bank. Corrie Vis (CentERdata) was very helpful in arranging the survey. Jansen thanks the ECB's Monetary Policy Research Division for its hospitality while working on this project. Any errors or omissions are our own responsibility. This paper presents the authors' personal opinions. Views expressed in this paper do not necessarily coincide with those of De Nederlandsche Bank, the European Central Bank, or the Eurosystem. Authors' e-mails: d.jansen@dnb.nl; c.a.b.van.der.cruijssen@dnb.nl; j.de.haan@dnb.nl.

1. Introduction

During the last few decades, many central banks have become increasingly more transparent, placing greater emphasis on external communications to financial markets and the general public (Dincer and Eichengreen 2007, 2009; Geraats 2009). Nowadays, the conventional wisdom is that independent central banks should be transparent and accountable, meaning that they have a responsibility to publicly explain their actions and the reasoning underlying those actions. In addition, transparency may enhance the effectiveness of monetary policymaking (Sibert 2009). According to the International Monetary Fund (2000), by “making the objectives of monetary policy public, the central bank enhances the public’s understanding of what it is seeking to achieve, and provides a context for articulating its own policy choices, thereby contributing to the effectiveness of monetary policy.”¹

Transparency requires that central banks communicate with the outside world. So far, the literature has mostly equated the outside world with financial market participants.² Blinder et al. (2008, p. 941) point out that “it may be time to pay some attention to communication with the general public.”³ In fact, we know very little about the general public’s knowledge of monetary policy, how information on central banks reaches the public, and how this information contributes to knowledge.⁴ More importantly, we have no clear understanding of whether knowledge of monetary policy is relevant for the public’s decision making.

¹Research suggests that central bank transparency enhances the effectiveness of monetary policy, although not all evidence points in the same direction (van der Cruysen and Eijffinger 2010b).

²There are many contributions to this literature, including Jansen and de Haan (2005), Ehrmann and Fratzscher (2007), Brand, Buncic, and Turunen (2010), Jansen (2011), Neuenkirch (2012), and Bulir, Čihák, and Jansen (2013, 2014).

³Similarly, Berger, Ehrmann, and Fratzscher (2006, p. 29) point out that “we have limited ourselves to an analysis of the transmission of communication, leaving open the question how communication is actually received by the final addressee, the general public. We leave this up to future research.”

⁴It should be noted that interest in this topic is growing; see, e.g., Hayo and Neuenkirch (2014). Another related paper is van der Cruysen and Eijffinger (2010a). However, they only measure knowledge about the European Central Bank’s disclosure practices.

Our paper contributes to the literature by charting what the general public knows about the objectives of the monetary policy of the European Central Bank (ECB). The ECB has specified its objective as follows: “The primary objective of the ECB’s monetary policy is to maintain price stability. The ECB aims at inflation rates of below, but close to, 2 percent over the medium term.”⁵ To study the general public’s knowledge, we conducted a survey of Dutch households. We presented participants with eleven statements about the ECB’s objective. Four of these statements were based on the ECB’s specification of its objective, while the remaining seven were false statements drafted by us. We asked respondents to answer whether they thought a particular statement was true or false.

Based on this survey, we address the following three questions. First, what does the general public know about the objectives of the ECB? Second, how does the general public obtain its information? Finally, to what extent is the public’s knowledge about the objectives of monetary policy relevant? To answer the third question, we focus on the relationship between knowledge and the respondents’ inflation expectations. The ability to make a reasonable assessment of future price developments is important for adequate household financial decision making (Katona 1975, Juster 1981, and Wärneryd and Wahlund 1985). Therefore, we assess whether knowledge about ECB objectives affects the ability to formulate realistic and accurate inflation expectations.

To organize our analysis, we use the framework proposed by Blinder and Krueger (2004) in their study on the determinants of opinions on U.S. economic policy. They find that ideology is the most consistently important determinant of opinions, while measures of self-interest are least important. Knowledge about the economy is important to explain positions on some policy issues, but not all. In contrast to Blinder and Krueger, we focus on knowledge rather than opinions.⁶

⁵Source: <http://www.ecb.int/mopo/html/index.en.html> (URL last accessed on November 19, 2014).

⁶Our work differs from the standard Eurobarometer survey carried out by the European Commission which also focuses on opinions. However, since 2007 Eurobarometer has conducted special surveys regarding the level of knowledge about economic indicators. On average, Europeans are unaware of economic growth rates, inflation rates, and unemployment rates, but do think that these figures are important for political decision making (European Commission 2010).

Our three main findings are as follows. First, we find that the respondents' knowledge about the ECB's policy objectives is far from perfect. The average number of correct answers to our eleven statements is less than five. There are interesting differences along demographic lines. We find that male respondents with strong self-interest, a clear ideology, and a high social status are more likely to have correct knowledge, while elderly respondents are more likely to have less knowledge about the ECB's monetary policy.

Second, the results indicate that respondents with a strong desire to be informed acquire more information on the ECB's policies. As a result, their knowledge is better than the knowledge of respondents with a relatively weak desire to be informed. Also, we find that respondents generally claim to have much better knowledge than they actually have. Turning to sources of information, our results also indicate that individuals who use the media more intensively to obtain information have better knowledge about monetary policy. This suggests that communication may be useful to improve the public's knowledge. It will be a challenge, however, to reach individuals who have no desire to be informed.

Finally, we find strong evidence that better knowledge about the objectives of ECB monetary policy is related to an individual's ability to form inflation expectations. This part of the paper links to the literatures on knowledge heterogeneity (e.g., Orphanides and Williams 2004), financial literacy (e.g., Lusardi and Mitchell 2007a, Bruine de Bruin et al. 2010), and household surveys of expectations (e.g., Souleles 2004). Using probit regressions, we find that individuals who answer more questions about ECB objectives correctly have a higher probability of having realistic inflation expectations. In addition, our results suggest that respondents' knowledge about ECB monetary policy is negatively related to their absolute inflation forecast errors. These results suggest there are important returns to this kind of knowledge, as the ability to make a reasonable assessment of future price developments is important for adequate household financial decision making.

2. Framework

We relate the level of understanding of monetary policy to a number of factors. First, the expected benefits from obtaining information

are important. If someone sees no benefits from being informed on monetary policy, she will have little incentive to obtain the relevant information, so that her level of understanding of central bank policies will be low. If someone does have an interest, the type of sources may become relevant. Most people do not receive their information on monetary policy directly from the central bank but via “intermediaries,” like television, radio, and newspapers. The role of these intermediaries in affecting the public’s understanding of central bank policies has been largely ignored in the literature. Our survey measures how and where respondents get their information on the ECB. Third, the quantity of information received through the various sources may play a role in someone’s understanding of monetary policymaking. Receiving more information may simply lead to a more thorough understanding of what central banks are doing. Finally, several other variables, such as the level of education, the type of job, or the level of income, may also be relevant.

The factors mentioned above are, in all likelihood, interrelated. To formalize the interdependencies between understanding (U), desire to be informed (D), sources of information (S), and quantity of information (Q), we use the framework as proposed by Blinder and Krueger (2004). We assume that the desire of person i to be informed is a function of self-interest (SI), ideology (ID), education (ED), and a vector X that contains control variables like sex and age.

$$D_i = f_1(SI_i, ID_i, ED_i, X_i) + e_{1,i} \quad (1)$$

As mentioned, one of the crucial reasons that someone may want to be informed about monetary policy is self-interest. Self-interest may arise for several reasons, and choosing the best measure for self-interest is not obvious (Blinder and Krueger 2004). Our preferred measure in the empirical analysis is income. Income may be relevant, as the real value of someone’s income will be affected by inflation. We consider two alternatives. First, an individual may have an interest because her line of work is closely influenced by monetary policy. Second, an individual may be interested in monetary policy because she has experienced negative consequences due to the financial and economic crisis.

Following Blinder and Krueger (2004), we further posit that

$$S_i = f_2(D_i, SI_i, ID_i, ED_i, X_i) + e_{2,i} \quad (2)$$

and

$$Q_i = f_3(D_i, SI_i, ID_i, ED_i, X_i) + e_{3,i}. \quad (3)$$

The sources of information and the quantity of information depend on the desire to be informed, self-interest, ideology, education, and various control variables. The more someone wants to be informed about monetary policy and the stronger the self-interest, the more (different) sources of information she will use and the more information she will acquire. Also, ideology may affect the sources of information used. People probably consult those sources of information that are closest to their own ideological position. Finally, the level of education is also likely to affect sources and quantity of information. First, highly educated persons will read different newspapers than people with low levels of education. Also, for these individuals newspapers may be more important than, for instance, television. Second, individuals with lower degrees of education may use fewer sources of information. Third, there is evidence that the level of education matters for the extent to which news is retrieved from various media sources. For instance, Grabe, Kamhawi, and Yegyan (2009) show that highly educated persons are relatively better at remembering items read in newspapers and on the Internet. In contrast, the lower educated perform best at retrieving news presented on television.

Following Blinder and Krueger (2004), we assume that understanding of monetary policymaking (U) is determined by ED , D , S , Q , and X , so that we can write

$$U_i = f_4(ED_i, D_i, S_i, Q_i, X_i) + e_{4,i}. \quad (4)$$

Section 5 will discuss various empirical models for U_i . In the analysis, we include the variables that we assume to affect D , S , and Q as well as these variables themselves. The reason for also including D , S , and Q is that the fit of our models explaining these variables

is rather poor. In our model, the coefficients of D , S , and Q reflect the impact of the unexplained variation of these variables on U . The next section describes our survey and explains how we use the respondents' replies to construct proxies for the variables used in the empirical analysis.

3. The Survey

In April 2009 we sent a questionnaire to members of the DNB Household Survey (DHS). The DHS, formerly known as the CenterER Savings Survey, is a panel study initiated in 1993 by CenterERdata, a research institute affiliated with Tilburg University. The survey comprises a representative sample of the Dutch population. Among other things, the panel provides data on income, housing, mortgages, loans, and personal characteristics.⁷ Our questionnaire was sent out via the Internet to 2,369 regular members of the DHS (sixteen years and older) from April 17 until April 21, 2009. Compared with surveys conducted by telephone or mail, the response rate to this continuous Internet-based survey is usually very high. In our case, the response rate is 70 percent, which corresponds to 1,659 individuals.

One potential concern regarding our data source is that participation in the DHS may have improved the knowledge of participants, as questions related to monetary policy have been submitted to panel participants before. For instance, van der Crujsen and Eijffinger (2010a) used the DHS to study transparency of the ECB. If participation improves knowledge, our sample may no longer be representative of the Dutch population at large. Therefore, we examined whether length of participation in the panel was related to the level of knowledge about monetary policy. We found no evidence of a significant relationship between these variables. Another potential concern is that participants are aware that the survey is affiliated with De Nederlandsche Bank (DNB) and therefore feel more pressured to provide correct answers. However, in practice, all contacts

⁷Information on the DHS is available at <http://www.centerdata.nl/en/projects-by-centerdata/dnb-household-survey-dhs> (URL last accessed on March 21, 2012). See also Teppa and Vis (2012). Work using the DHS includes Hurd, van Rooij, and Winter (2011), van Rooij, Lusardi, and Alessie (2011, 2012), and Jansen, Mosch, and van der Crujsen (2015).

with the survey participants are handled by CentERdata, and it is not mentioned in our questionnaire that DNB has commissioned it. So, it seems more likely that participants associate the survey with the University of Tilburg rather than DNB.

Appendix 1 lists our survey questions. The first three questions are general in nature. First, we asked respondents to indicate to what extent they were affected by the ongoing economic and financial crisis. We use this information to construct a proxy for self-interest which ranges from 1 (not hit by the crisis) to 4 (very much hit by the crisis). Second, we asked whether respondents had occupations related to economic, monetary, or financial issues. The responses to this question are used to construct another self-interest variable that ranges from 1 to 5. When the respondent's work is not related to economic, monetary, or financial issues, this variable equals 1, while it is 5 when the respondent's job relates to these matters on a daily basis.

Third, we asked respondents to identify their political orientation. The answers to this question were used to construct a *no ideology* (*ID*) dummy, which is 1 for respondents indicating not to have thought much about politics. The motivation for this *ID* variable is Blinder and Krueger's (2004) finding that people without a clear ideology have less knowledge about economic policy.

The survey continued with a short introduction of the ECB. As we are interested in knowledge about monetary policy, we only informed respondents that "the ECB is the central bank for Europe's single currency, the euro. Since 1999 the euro has been introduced in 16 European countries." Next, we asked respondents to rate their knowledge about the ECB and economic policy in general. We added this question to assess whether there is a mismatch between an individual's perceived knowledge about the ECB and her actual knowledge. Central bank communication policy will have greater difficulty in reaching those people who incorrectly believe they are well informed.

The subsequent question was designed to measure explicitly the respondents' desire to be informed (*D*). We asked the following question: "How important is it to you to keep well informed about the policies of the European Central Bank (ECB)?" There were five possible answers ranging from "not important at all" to "extremely important," which are ranked on a scale between 1 (low) and 5

(high). In question 6 we asked why respondents want to be informed about ECB policy. Respondents could indicate the importance they attach to a number of reasons.

Question 7 is arguably the most important in our survey, as it measures the respondents' knowledge of monetary policy. It consists of a list of eleven statements on the main objective of the ECB's monetary policy. Four of these statements were based on the ECB's specification of its objective, while the remaining seven were false statements drafted by us. For each statement, respondents were asked to answer whether the statement is true or false. As proxy for knowledge (U) we use the number of correct answers. We also analyze whether there are differences between the fraction of correct answers to true and false statements.

In the questionnaire, we included a portion of text stressing that there was no need to search for the correct answers (see appendix 1). We even explicitly mentioned that participants need not worry about giving an incorrect answer. By including these comments, we wanted to minimize the likelihood that people used Internet sources (such as the ECB website) to search for information while completing the survey. Of course, we cannot exclude that people searched for correct answers. Still, searching for the answers to the eleven statements would have taken quite some time. Also, we did not offer participants any monetary incentives for answering more questions correctly and survey responses are anonymous, so that it is not possible to check which individual gave how many correct answers. Therefore, it seems unlikely that a significant portion of the respondents engaged in searching behavior. Moreover, if people did search for information, this would only strengthen our finding that knowledge on ECB monetary policy is far from complete.

The remaining questions of our survey referred to the sources of information on the ECB. We first asked which sources of information respondents use to get information about the policies of the ECB. The respondents could indicate whether they used the various sources—such as television, radio, or Internet—regularly, occasionally, or never. We use the answers to measure the use of information on a scale between 1 (never) and 3 (regularly). These measures are proxies for Q (quantity of information).

Along the lines of Blinder and Krueger (2004, p. 341), we construct two measures for the *intensity* and *lack of diversity* of

information. Let s_1 , s_2 , and s_3 be the number of sources through which respondents “never,” “occasionally,” or “regularly” obtain information on the ECB, respectively, and let s denote the total number of sources (s can be less than six because of non-response). Then QH (“quantity high”) and QL (“quantity low”) are defined as s_3/s and s_1/s , respectively. For example, if a respondent has a QH of 0.25 and a QL of 0.5, this respondent reads or hears regularly about the ECB via 25 percent of the sources, but via half of the sources she never gets information on the ECB.⁸

The final question is: “Which of the following is your most important source of information on the policy of the European Central Bank?” We use the response to this question to construct a dummy that measures whether or not a particular medium is the most important source of information. These dummies are proxies for the variable S in our model.

We also have detailed background information on the DHS panel members, which we use to construct various control variables (X): a dummy that is 1 if the respondent is male (*male*), the age of the respondent (*age*), and a dummy that is 1 if the survey participant is living together with a partner (*partner*). As to the latter, it seems plausible that respondents with a partner are able to draw from a larger pool of information. Furthermore, we control for the degree of urbanization by including a variable called *city*, which ranges from 1 (respondent lives in a rural area) to 5 (respondent lives in a very strongly urbanized area). We also have information on the size of households.⁹ In addition, we control for the social-economic status of the respondent (*status*). This variable ranges from 1 (weak status) to 5 (strong status). This variable—originally defined by Statistics Netherlands—takes a person’s profession into account and whether she holds a managing position. If she holds such a position, the number of employees is also taken into account. We also use DHS background information to create an education dummy *ED* that

⁸There is a subtle difference regarding Blinder and Krueger’s measure for QL and ours. In Blinder and Krueger (2004), QL is based on sources which are used “rarely or never,” whereas in our questionnaire, QL is based on sources which are never used.

⁹We do not use household size in our regressions, as it did not turn out to be relevant. We do use the information on household size to study whether our sample is representative of the Dutch population.

is 1 for respondents who successfully completed higher vocational education and/or university education, and 0 otherwise. Finally, we use data on household income to construct our third self-interest variable, which ranges from 1 to 12, based on twelve monthly gross household income categories. The income variable (labeled *SI1*) is our preferred proxy for self-interest.¹⁰ The intuition is straightforward: high-income respondents would stand to lose more from high levels of inflation, so monetary policy would be highly relevant to them. *SI1* is 1 for respondents reporting a household income of €500 or less, and 12 when monthly household income is €7,500 or more.

4. Overview

Table 1 provides information on our respondents' gender, age, size of their household, gross monthly income, education level, where they live, and whether they are living with a partner. The average respondent turns out to be male, in his early fifties, living with a partner, and earning a gross monthly income of around €3,800. Table 1 also compares the survey means with the Dutch population based on data provided by Statistics Netherlands (CBS). On a number of dimensions, such as age, gender, and education, the sample is not fully representative for the Dutch population. We checked, therefore, whether reweighting observations would change our conclusions. This turned out not to be the case. Appendix 2 provides more details on the effect of weighting.

Figures 1–3 give information about the extent to which respondents were hit by the financial crisis, their work environment, and their ideological position. Figure 1 shows that about 45 percent of the respondents answered that they were either very much or somewhat affected by the crisis. Figure 2 reveals that about 70 percent of the respondents are not professionally interested in monetary policymaking, thereby ensuring that the survey captures the understanding of the “public at large.” Finally, figure 3

¹⁰Regression results using the other two proxies of self-interest (“affected by crisis” and “work related to economic issues”) to explain knowledge are available in appendix 2 of van der Cruijssen, Jansen, and de Haan (2010, p. 43). From that analysis, we concluded that income is more closely linked to knowledge than the alternative measures for self-interest.

Table 1. Summary Statistics

	Survey Mean	Population Mean
Sex (percent males)	55.1%***	49.0%
Age	52.6***	46.9
Household Size	2.5***	2.2
Partner (1 = living together with partner, 0 = otherwise)	0.78***	0.62
City Weighted Average (1 = rural, 5 = highly urbanized)	2.99	3.03
Monthly Gross Household Income (in €1,000)	3.77***	4.68
Education (0 = low, 1 = high)	0.36***	0.25
Region	Proportion Differences Range from −0.03 to 0.01	

Source for Population Means: Statistics Netherlands (CBS) data for 2009.

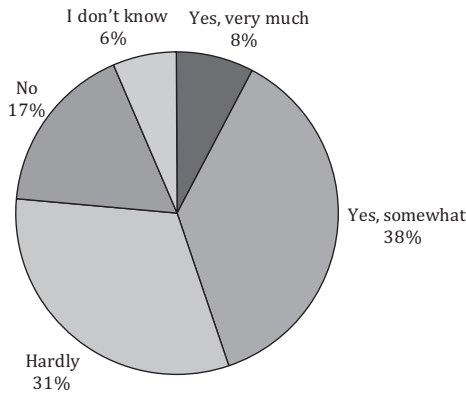
Notes: For variables sex and age, we are able to use specific data on the population over age sixteen. For education, CBS data cover ages fifteen to sixty-five. For other variables, we use data on the total population. For income, data were only available yet for 2008. Education is coded as 1 if higher vocational education and/or university education was the highest degree, and 0 if otherwise. For region, we use the proportion of survey participants living in province x minus the actual proportion of people aged fifteen years or older in province x. *** denotes significant differences at the 1 percent level between survey and population means.

presents information about the self-declared ideological position of the respondents. About 20 percent of the respondents indicate that they have not thought much about their political position.

Table 2 shows that almost 63 percent of the respondents indicate that being informed about ECB policies is either somewhat important or very important. When asked for the reasons they want to be informed, 57 percent of the respondents consider the importance of the ECB for the economy as the crucial reason to be informed about the ECB (see table 3). A majority of the respondents report that they “just like to be informed.” Table 3 also shows that (i) 52 percent and (ii) 66 percent of respondents consider it a somewhat/very/extremely important reason to be informed because of the effect of ECB policy on (i) “my personal or family income” and (ii) “how much my money can buy.”

Figure 1. Impact of Financial Crisis on Respondents

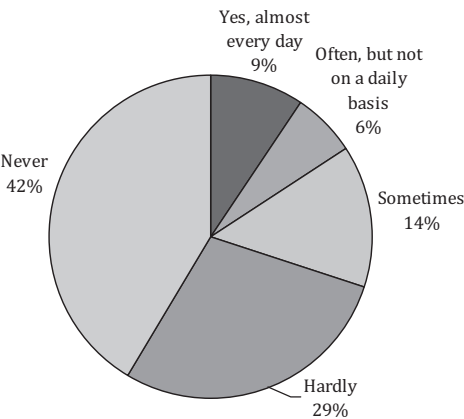
Q1. Has the current economic and financial crisis had negative consequences for you, or do you expect such consequences in the near future?



Notes: Based on replies by 1,659 respondents.

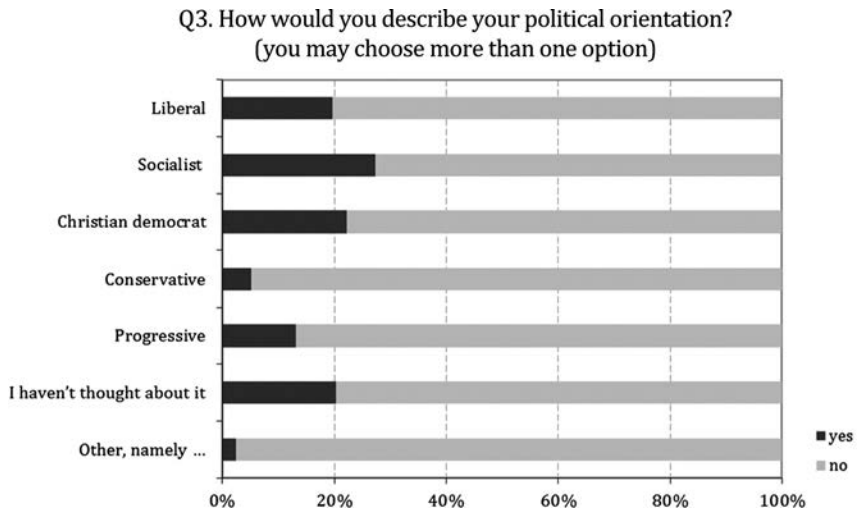
Figure 2. Working Environment of Respondents

Q2. Does (or did) your work relate to economic, monetary or financial issues?



Notes: Based on replies by 1,659 respondents.

Figure 3. Self-Declared Ideological Position



Notes: Based on replies by 1,659 respondents.

Table 2. Importance of Being Informed on European Central Bank

	Frequency	Percentage	Cumulative
Not Important at All	71	4.3	4.3
Not Very Important	366	22.1	26.3
Somewhat Important	797	48.0	74.4
Very Important	247	14.9	89.3
Extremely Important	58	3.5	92.8
I Don't Know	120	7.2	100.0

Notes: N = 1,659. This table shows the distribution of answers to question 5 of our questionnaire: “How important is it to you to be well informed on the policies of the European Central Bank (ECB)?”

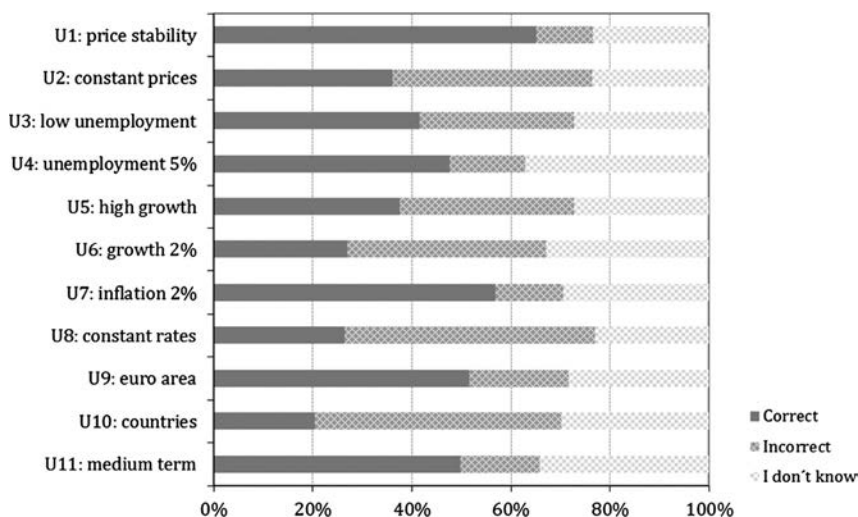
Figure 4 shows the distribution of correct and incorrect answers to the eleven statements about the ECB’s objectives (question 7 in appendix 1). The statement that the main objective of the ECB is price stability received by far the highest percentage of correct

Table 3. Reasons Why Respondents Want To Be Informed

	(a) Not Important at All	(b) Not Very Important	(c) Somewhat Important	(d) Very Important	(e) Extremely Important	(f) I Don't Know	Rank (Based on c+d+e or d+e)
ECB policy affects my personal or family income.	9.0%	18.9%	32.4%	16.0%	4.0%	19.8%	4
ECB policy influences how much my money can buy.	4.8%	12.3%	34.3%	26.0%	5.5%	17.0%	3
ECB policy affects my business/job/profession.	26.9%	19.1%	20.1%	10.0%	2.7%	21.2%	6
ECB policy affects the value of my stocks or other investments.	26.4%	16.0%	20.5%	14.0%	4.1%	19.0%	5
ECB policy is important for the economy.	1.8%	4.8%	21.9%	38.8%	18.1%	14.6%	1
I just like to keep informed well.	5.7%	14.4%	32.4%	25.1%	11.0%	11.3%	2

Notes: N = 1,468.

Figure 4. Understanding: Distribution of Answers Per Question



Notes: N = 1,659. The horizontal bars denote the percentage of correct (darkest gray) and incorrect (crossed dark gray) answers to the eleven statements on the ECB's main objective (see question 7 in appendix 1). The checkered light gray bars denote the percentage of respondents who answered "I don't know."

answers (65 percent). Details about the inflation target are less well known. Importantly, few respondents are aware that the ECB does not define its objective in terms of inflation in each individual euro-area country. The (false) statement that the ECB's objective applies to all euro-area countries separately (statement 10) received the lowest score of correct answers (20 percent).¹¹ In fact, of the 1,165 respondents who did not opt for "I don't know," only 29 percent gave the correct answer. One interpretation is that the public is actually misinformed rather than uninformed on this issue.¹² This

¹¹In contrast, van der Cruijssen and Demertzis (2011) show that professionals do realize that the ECB goal does not apply to national inflation rates.

¹²We tested, per question, whether more than 50 percent of the respondents who do answer the true/false statements give the correct answer. If not, one could argue that the public at large is misinformed rather than uninformed. Our results (available on request) suggest that misinformation would apply to statements 6,

finding is potentially worrisome. It implies that even when the ECB meets its inflation goal, it can suffer a loss of credibility in countries in the euro area with an inflation rate above 2 percent. As a result of insufficient knowledge, it is likely that an important share of the public at large will interpret deviations of national inflation from “close to but below two percent” as a shortcoming of ECB policy. The statements on economic growth (statement 6) and constant rates (statement 8) also show a low number of correct answers. In these cases, only 40 percent (statement 6) and 34 percent (statement 8) of the respondents who believed to know the answer submitted correct answers.

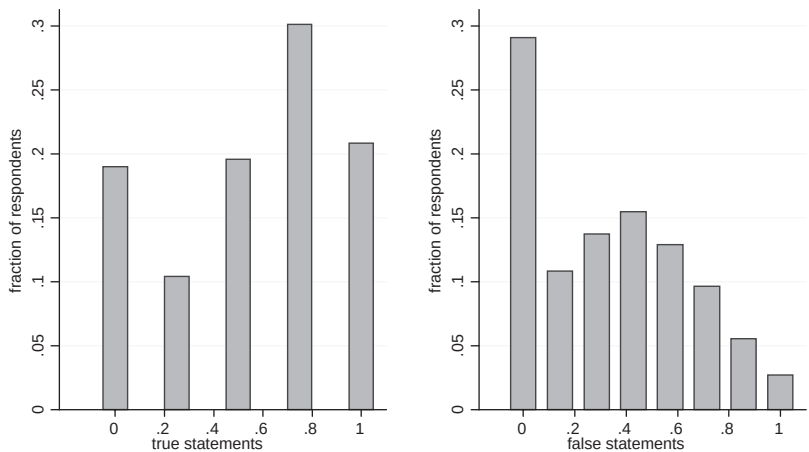
Finally, the difference between true and false statements turns out to be important. The true statements were statements 1 (objective is price stability), 7 (inflation close to but below 2 percent), 9 (objective applies to the euro-area average), and 11 (objective applies to the medium term) of question 7 of our survey. The horizontal axis of figure 5 shows the fraction of correctly answered true (left panel) and false (right panel) statements. The vertical axis shows the fraction of respondents per category.¹³ Clearly, respondents did much better at answering true statements. This suggests that when in doubt about the correct answer, respondents chose “true” as their default answer. As a consequence, we may be over-estimating the level of knowledge about monetary policy. For instance, we were surprised by the number of people (50 percent) who answered statement 11 (the ECB’s objective applies to the medium term) correctly, as we considered this to be one of the more difficult statements to answer. However, if respondents chose “true” when in doubt, the estimate of 50 percent would be optimistic.

Next, we relate respondents’ self-declared knowledge to their actual knowledge. Figure 6 shows that there is a positive link

8, and 10, where the fraction of correct answers is well below 0.5. Note, however, that these questions were all false statements, which were harder for participants to answer in general.

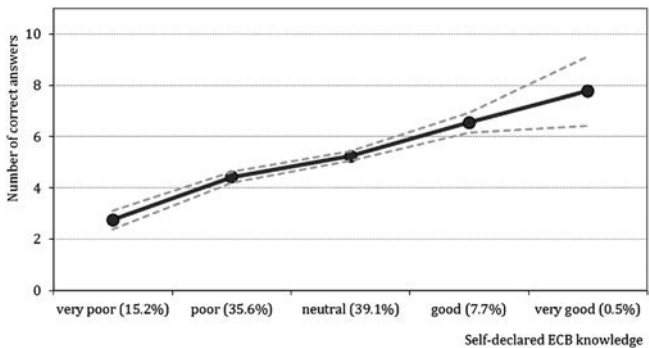
¹³ As there were only four (seven) statements which were true (false), the distribution of the fractions of correct answers can only take a limited number of values.

Figure 5. Distribution of Correct Answers to True and False Statements



Notes: The histograms show on the horizontal axis the proportion of correct answers to true (left panel) and false (right panel) statements. The vertical axis measures the fraction of respondents per category. True statements were statements 1, 7, 9, and 11 of question 7 of our survey (see appendix 1). The remaining seven statements were false. Histogram is based on answers by 1,659 respondents.

Figure 6. Actual vs. Self-Declared Knowledge



Notes: N = 1,628. Response shares are shown in parentheses (1.9 percent of respondents did not report self-assessment). The circles denote the average number of correct questions; dotted lines denote the 95 percent confidence interval.

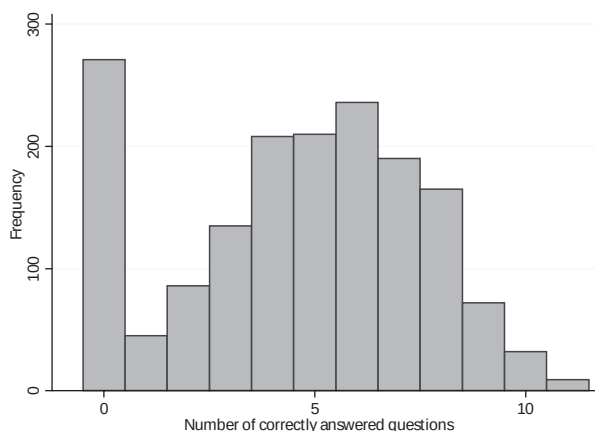
Table 4. Most Important Source of Information on the ECB

	Proportion	Std. Err.	95% Confidence Interval
Television	0.42	0.01	(0.39,0.44)
Radio	0.05	0.01	(0.04,0.06)
Newspapers	0.33	0.01	(0.30,0.35)
Magazines	0.02	0.00	(0.01,0.02)
Internet	0.05	0.01	(0.04,0.06)
Friends/Relatives/ Colleagues	0.01	0.00	(0.00,0.01)
I Don't Use Any Sources on ECB Policy	0.08	0.01	(0.07,0.10)
I Don't Know	0.04	0.01	(0.03,0.05)
Other	0.01	0.00	(0.00,0.01)
Notes: N = 1,400.			

between the two; on average individuals with higher self-declared ECB knowledge answer more questions correctly (correlation: 0.36). However, almost all survey participants who indicate having good or very good knowledge (8.3 percent of the respondents) “fail” the test, as they do not *fully* understand the ECB’s main objective. The average number of correct answers is 6.5 for the “good” category, and 7.8 for the “very good” category. Especially in this last category, the dispersion is quite large. Although central bank communication can be used as a tool to improve knowledge, it will be difficult to reach those individuals who are not aware of these knowledge gaps. Furthermore, overconfident individuals may fail to prepare optimally for the future (Marzilli Ericson 2011).

Finally, table 4 shows the answers to the question concerning the most important source of information. It is clear that television is by far the most important source of ECB information. Almost 42 percent of the respondents put this intermediary on top, followed by newspapers (almost 33 percent). The other sources of information receive low scores.

Figure 7. Understanding: Distribution of Number of Correct Answers



Notes: N = 1,659. The histogram shows the number of correct answers to our eleven statements regarding the ECB's main objective of price stability.

5. Determinants of Knowledge about the ECB's Objectives

We now turn to explaining the respondents' knowledge about the ECB's monetary policy. As the dependent variable, we first use the number of correctly answered questions on monetary policy objectives. Following Blinder and Krueger (2004), zero scores were assigned if respondents answered "I don't know" to the knowledge questions. Figure 7 shows the distribution of the number of correct answers. Given the discrete and limited nature of this variable, we use models for analyzing count data.

The workhorse model in this field is the Poisson regression.¹⁴ This model assumes that the dependent variable y_i is independently Poisson distributed with density

$$f(y_i|x_i) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!} \quad (5)$$

¹⁴The following technical discussion draws on Greene (2003), Cameron and Trivedi (2005, 2007), and Winkelmann (2008).

and with mean parameter

$$E[y_i|x_i] = \mu_i = \exp(x_i'\beta), \quad (6)$$

where x_i is a vector of independent variables and β is a $k \times 1$ parameter. The model is estimated by maximum likelihood. One crucial assumption of the Poisson model is that the conditional mean equals the conditional variance. This assumption of *equidispersion* is often violated, as the data are either over-dispersed (when the variance is larger than the mean) or under-dispersed (when the variance is smaller than the mean). A violation of the assumption of equidispersion will affect the estimated standard errors, in which case it is necessary to adjust the regression model accordingly. One straightforward way to check for equidispersion is to compare the sample mean and variance of the dependent variable.

In our case, the standard Poisson assumption of equality between the conditional mean and the conditional variance is violated due to a clustering of the observations around zero. Of the 1,659 observations, 16.3 percent has a value equal to zero. First, the zero category includes those 0.8 percent of the respondents who either answered all knowledge questions incorrectly or who answered some questions but did not provide correct answers. Second, the zero category consists of those 15.6 percent of the respondents who answered “I don’t know” to all of the knowledge questions. Although these individuals may have had knowledge about monetary policy, they declined to answer any question. So, for each knowledge question, the respondent had to make two decisions. First, the respondent had to decide whether she wanted to answer the question. Second, if “I don’t know” was not selected, the respondent had to choose between “true” and “false.”

Therefore, it is important to model the decision to answer the knowledge questions in addition to analyzing the number of correctly answered questions. The literature on count data suggests two models: the hurdle model and the zero-inflated model. Winkelmann (2008, p. 189) recommends using a zero-inflated model once there is more than one reason for the occurrence of zeros. Therefore, our

preferred model is a zero-inflated Poisson (ZIP) model (Lambert 1992).¹⁵ This means the combined estimation of a logit regression for the probability of a zero outcome and a Poisson regression for the number of correct answers to question 7. The density for the ZIP model is (Cameron and Trivedi 2005, p. 681):

$$\begin{aligned} g(y) &= h(0) + (1 - h(0)) * f(0) & \text{if } y = 0, \\ &(1 - h(0)) * f(y) & \text{if } y \geq 1, \end{aligned} \quad (7)$$

where $h(.)$ denotes the logit density, $f(.)$ is the Poisson density as described above in equation (5), and y is the number of correct answers to question 7 of the survey. The explanatory variables in the two parts of the model can be different. However, in practice, the independent variables used in the logit regression are often also used in the Poisson part of the model (Cameron and Trivedi 2007, p. 127).

Table 5 presents the estimation results for four zero-inflated Poisson models, while table 6 shows marginal effects computed at sample means. Column 1 of table 5 presents the results if the variables in the X vector are included, as well as the desire to be informed (D), self-interest (SI), the no-ideology variable (ID), and education (ED). Respondents with a strong desire to be informed, strong self-interest, a high social status, and who live in urbanized areas are less likely to have a zero score in the logit model, i.e., they are more inclined to answer the knowledge questions. In contrast, respondents without a clear ideology are less likely to receive a positive score. The results for the Poisson regression suggest that male respondents with strong self-interest, a clear ideology, higher education, and a high social status are more likely to answer the statements about the ECB's monetary policy correctly. Elderly respondents are more likely to have less knowledge about the ECB's monetary policy. In contrast to our

¹⁵We compared the outcomes against a hurdle model (Mullahy 1986). The parameter estimates for the two models were very similar. This is probably due to the small fraction (0.8 percent) of “true” zeros. As noted by a referee, it is also possible to distinguish between “true” and “excess” zeros, and estimate separately a Poisson model for the subsample of respondents who answer knowledge questions, and a logit regression explaining the decision to decline to answer any questions, while excluding the “true” zeros. This alternative estimation strategy also leaves the conclusions unchanged. Results for both exercises are available upon request.

Table 5. Determinants of Knowledge about ECB Monetary Policy

	(1)		(2) S		(3) Q		(4) QH and QL	
	Poisson	Logit	Poisson	Logit	Poisson	Logit	Poisson	Logit
Desire to be Informed (ID)	0.01 (0.01)	-0.80*** (0.15)	-0.01 (0.02)	-0.57*** (0.21)	-0.01 (0.02)	-0.68*** (0.19)	-0.01 (0.02)	-0.63*** (0.18)
Self-Interest (SI1)	0.02*** (0.00)	-0.13*** (0.05)	0.02*** (0.00)	-0.19*** (0.06)	0.02*** (0.00)	-0.10* (0.05)	0.02*** (0.00)	-0.12*** (0.05)
No Ideology (ID)	-0.12*** (0.04)	0.86*** (0.21)	-0.11*** (0.04)	0.67*** (0.29)	-0.10*** (0.04)	0.52*** (0.25)	-0.11*** (0.04)	0.71*** (0.24)
Education (ED)	0.07*** (0.03)	0.43* (0.24)	0.05* (0.03)	0.27 (0.32)	0.07*** (0.03)	0.22 (0.28)	0.07*** (0.03)	0.33 (0.26)
Age	-0.00*** (0.00)	-0.00 (0.01)	-0.00*** (0.00)	0.00 (0.01)	-0.00*** (0.00)	0.00 (0.01)	-0.00*** (0.00)	0.00 (0.01)
Male	0.12*** (0.03)	-0.15 (0.21)	0.11*** (0.03)	0.20 (0.28)	0.10*** (0.03)	0.04 (0.24)	0.11*** (0.03)	-0.01 (0.22)
Partner	-0.07*** (0.03)	0.28 (0.26)	-0.08*** (0.03)	0.29 (0.35)	-0.07*** (0.03)	0.35 (0.30)	-0.08*** (0.02)	0.28 (0.26)
Status	0.02*** (0.01)	-0.38*** (0.10)	0.02*** (0.01)	-0.26*** (0.13)	0.02* (0.01)	-0.31*** (0.12)	0.02* (0.01)	-0.32*** (0.10)
City	0.01 (0.01)	-0.18** (0.07)	0.01 (0.01)	-0.11 (0.09)	0.01 (0.01)	-0.12 (0.09)	0.01 (0.01)	-0.15* (0.08)

(continued)

Table 6. Marginal Effects Based on Zero-Inflated Poisson Models

	(1)	(2) S	(3) Q	(4) QH and QL
Desire to be Informed (D)	0.36*** (0.08)	0.09 (0.09)	0.14 (0.09)	0.14 (0.09)
Self-Interest (SI1)	0.14*** (0.03)	0.14*** (0.03)	0.12*** (0.03)	0.13*** (0.03)
No Ideology (ID)	-0.93*** (0.20)	-0.75*** (0.22)	-0.67*** (0.20)	-0.79*** (0.20)
Education (ED)	0.17 (0.15)	0.22 (0.15)	0.30* (0.15)	0.23 (0.15)
Age	-0.02*** (0.00)	-0.02*** (0.01)	-0.02*** (0.00)	-0.02*** (0.00)
Male	0.65*** (0.15)	0.53*** (0.15)	0.53*** (0.15)	0.56*** (0.15)
Partner	-0.45*** (0.16)	-0.52*** (0.15)	-0.44*** (0.15)	-0.48*** (0.15)
Status	0.27*** (0.07)	0.18*** (0.07)	0.20*** (0.07)	0.21*** (0.07)
City	0.12** (0.05)	0.06 (0.05)	0.09* (0.05)	0.09** (0.05)
S or Q:				
TV		0.52** (0.15)	0.27** (0.13)	
Radio			0.16 (0.11)	
Newspapers		0.80*** (0.27)	0.39*** (0.12)	
Magazines			0.05 (0.14)	
Internet			0.01 (0.12)	
Friends and Relatives			-0.09 (0.16)	
Other		1.08*** (0.30)		
QH				1.40*** (0.49)
QL				-0.53* (0.30)

Notes: Robust standard errors are shown in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. The table reports marginal effects evaluated at sample means for the four zero-inflated Poisson models shown in table 5.

expectations, the same finding emerges for the partner variable. The marginal effects (shown in table 6) indicate that the included covariates have a significant impact on the number of correct answers, except for education.¹⁶ Although the desire to be informed is not significant in the Poisson part of the model, the overall marginal effect is significant. The intuition is that the desire to be informed has a strong effect on the decision to answer the question, as shown in the logit regression.

In column 2 of table 5, we add three variables reflecting the most important source of information on the ECB's monetary policy according to the respondents. The marginal effects in table 6 are significant for all three media variables. This shows that respondents who use information have better knowledge than those 8 percent of the respondents who do not use any information at all. Next, we replace S with Q , our proxy for the quantity of information. As it turns out, this proxy is positive and significant for the case of television and newspapers (table 6, column 3). This means that greater use of television and newspapers to learn about ECB policies has a positive relationship with actual knowledge about ECB objectives. Finally, we insert the intensity of information variable QH and lack of diversity variable QL (column 4 of table 5). Respondents with relatively low information diversity—a relatively high share of sources through which they *never* obtain information on the ECB—are less inclined to answer the knowledge questions. Given that respondents answer the knowledge questions, those with high information intensity—a relatively high share of sources through which they *regularly* receive information on the ECB's policy—have a better understanding of the ECB's goal. Overall, both QH and QL have a significant effect on the number of correct answers (table 6, column 4). As expected, QH is positively and QL is negatively related to knowledge.

Overall, the results for the three sets of media variables suggest that obtaining more information leads to better knowledge about monetary policy.¹⁷ This implies that the media channel may be a useful channel for central banks to improve the public's knowledge.

¹⁶Note that education may still play a significant role via the desire to be informed (for further discussion, see van der Cruysen, Jansen, and de Haan 2010).

To assess whether the difference between true and false statements is important, we repeated the analysis for determinants of knowledge. As we now use models where the dependent variables are the fraction of correctly answered true and false statements, we revert to the quasi-likelihood method proposed by Papke and Woolridge (1996). Once again, we estimate three types of models, based on the type of media variables (S , Q , or QH and QL) included. Table 7 shows the estimation results. The finding that obtaining information through the media improves knowledge continues to hold for true and false statements. There is even a clearer role for QL , but only in the case of the true statements. Still, some interesting differences compared with the results based on the total number of correct answers show up. First, the desire to be informed has a significant effect in all three models, but only for the true statements. Second, the demographic variables continue to be important, but often only in the case of the false statements.

6. Relationship Between Knowledge and Inflation Expectations

We now turn to the final question: To what extent are inflation expectations and knowledge about the ECB's objective related? To address this issue, we use the following standard question from the DHS: "What do you think is the most likely price increase (increase of consumer prices) over the next twelve months?" Figure 8 shows the distribution of the answers to this question. The expectations range between 1 and 10 percent, with the mode and median at 2 percent, while the mean is around 2.7 percent.¹⁸

Our analysis relates to three streams of research. First, several papers document that survey data on inflation expectations are inconsistent with the rational expectations hypothesis (Blanchflower

¹⁷As suggested by one of the referees, we have also estimated a Heckman selection model. The results (available on request) mirror the estimations in table 5.

¹⁸Admittedly, our test is indirect, as these expectations pertain to Dutch consumer prices. Expectations for the euro-area inflation are, however, not included in the DHS. It should also be noted that the participants were asked to submit their inflation expectations in a particular form: only natural numbers between 1 percent and 10 percent were possible.

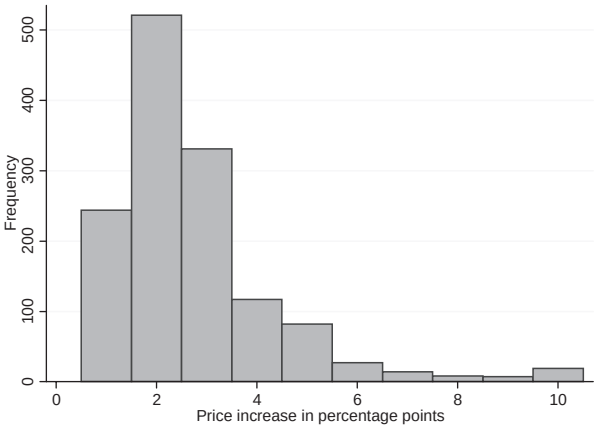
Table 7. Determinants of Knowledge: Difference between True and False Statements

	S		Q		QH and QL	
	(1) True	(2) False	(3) True	(4) False	(5) True	(6) False
Desire to be Informed (D)	0.25*** (0.05)	-0.05 (0.04)	0.24*** (0.06)	0.00 (0.05)	0.26*** (0.05)	-0.00 (0.05)
Self-Interest (SI1)	0.05*** (0.02)	0.06*** (0.02)	0.05*** (0.02)	0.05*** (0.02)	0.05*** (0.02)	0.06*** (0.01)
No Ideology (ID)	-0.48*** (0.12)	-0.25** (0.11)	-0.49*** (0.11)	-0.21* (0.11)	-0.56*** (0.11)	-0.27*** (0.10)
Education (ED)	-0.02 (0.09)	0.13 (0.08)	-0.02 (0.09)	0.18** (0.08)	-0.03 (0.09)	0.14* (0.08)
Age	-0.01* (0.00)	-0.01*** (0.00)	-0.01*** (0.00)	-0.01*** (0.00)	-0.01*** (0.00)	-0.01*** (0.00)
Male	0.05 (0.09)	0.26*** (0.08)	0.00 (0.09)	0.28*** (0.08)	0.04 (0.09)	0.29*** (0.08)
Partner	-0.10 (0.09)	-0.29*** (0.08)	-0.06 (0.09)	-0.26*** (0.08)	-0.06 (0.09)	-0.28*** (0.08)
Status	0.03 (0.04)	0.12*** (0.04)	0.05 (0.04)	0.12*** (0.04)	0.05 (0.04)	0.13*** (0.04)
City	0.03 (0.03)	0.03 (0.03)	0.05* (0.03)	0.05* (0.03)	0.05* (0.03)	0.05* (0.03)

(continued)

Table 7. (Continued)

Figure 8. Inflation Expectations of Survey Participants



Notes: This graph shows the distribution of inflation expectations for the participants in our survey who answered all knowledge questions. Of the 1,659 respondents to our survey, 1,370 answered the regular DHS question on inflation expectations. The precise wording of the question is: “What do you think is the most likely price increase (increase of consumer prices) over the next twelve months?” The responses were given in April 2009.

and Kelly 2008). Only a small portion of respondents form expectations consistent with rationality. Possible reasons for this heterogeneity include the use of different information sets by agents and different capacities of agents for processing information. In the framework of Carroll (2003), households obtain macroeconomic news using media sources. In each period, individuals absorb this news with a certain probability. Therefore, it may take time for news of changed macroeconomic circumstances to reach the entire population. Carroll finds that differences between household expectations and the views of professional forecasters narrow when inflation is more significant, probably because of increased media coverage and higher household interest. In learning models, individuals make statistical inferences about the unknown parameters governing the evolution of the economy. Orphanides and Williams (2005) find that the presence of learning increases the sensitivity of inflation expectations to shocks. From the perspective of the present paper, the result that inflation expectations under learning are much less sensitive to

inflation when the objectives of the central bank are known by the public is highly relevant. This finding suggests that better knowledge about the central bank may lead to more accurate expectations.

Second, also in the financial literacy literature the role of knowledge is crucial. For instance, lack of knowledge causes households not to plan for retirement (Lusardi and Mitchell 2007b) and to borrow at higher interest rates (Lusardi and Tufano 2009). Using RAND's American Life Panel, whose members were recruited from respondents participating in the Michigan Survey of Consumers in 2007, Bruine de Bruin et al. (2010) report that individuals with lower financial literacy have higher inflation expectations. Van Rooij, Lusardi, and Alessie (2011) report that Dutch households also exhibit fairly low levels of financial knowledge. Financial illiteracy is not only widespread, but is particularly severe in certain demographic groups. Especially the elderly and women display very low levels of economic knowledge.

Third, there is a well-established literature regarding survey data on household expectations.¹⁹ In an early contribution, Jonung (1981) shows how perceived and expected rates of inflation in a cross-section of Swedish households differ between several demographic groups. Differences between demographic groups are also found for the United States by Bryan and Venkatu (2001). Using the household data underlying the Michigan Index of Consumer Sentiment, Souleles (2004) finds evidence that household expectations are biased as well as inefficient. However, expectations are still economically meaningful, as they are useful in forecasting future consumption. Finally, Branch (2004, 2007) uses micro data from the Michigan Survey of Consumers over the period 1977 to 1993 to study endogenous model uncertainty and models incorporating sticky information as introduced by Mankiw and Reis (2002).

We now turn to the findings. Table 8 gives a breakdown of inflation expectations reported by DHS members in April 2009 along various dimensions. In line with the literature, we find that inflation expectations vary across demographic groups. In particular, men have lower inflation expectations than women ($p < 0.05$) and

¹⁹Most papers focus on aggregate household expectations. For expectations of experts, there is more research on individual forecasts. See, for instance, Capistrán and Timmermann (2009) or Dovern, Fritsche, and Slaálek (2012).

Table 8. Summary Statistics for Inflation Expectations

	Mean	Standard Deviation	Number of Observations
Full Sample	2.73	1.64	1,370
Breakdown By:			
Age			
Younger than 20	3.12	1.78	33
Between 20 and 39	2.61	1.56	274
Between 40 and 64	2.69	1.60	719
Between 65 and 79	2.83	1.75	316
80 or Older	3.04	2.00	28
Gender**			
Male	2.64	1.53	771
Female	2.83	1.77	599
Partner			
Living with a Partner	2.74	1.66	1,061
Not Living with a Partner	2.68	1.59	309
Income			
Low Income	2.89	1.81	605
Medium Income	2.63	1.46	601
High Income	2.48	1.55	164
Education**			
No Higher Vocational Education or University	2.81	1.72	874
Higher Vocational Education or University	2.58	1.48	491
Social-Economic Status			
1 (= Lowest)	4.32	2.58	19
2	2.91	1.76	297
3	2.77	1.70	333
4	2.59	1.40	441
5 (= Highest)	2.57	1.64	276
City			
1 (= Least Urbanized)	2.83	1.87	224
2	2.56	1.38	311
3	2.67	1.52	289
4	2.94	1.92	324
5 (= Most Urbanized)	2.60	1.43	215

(continued)

Table 8. (Continued)

	Mean	Standard Deviation	Number of Observations
Desire to be Informed			
1 (= Lowest)	3.20	2.01	59
2	2.86	1.74	289
3	2.63	1.54	677
4	2.63	1.56	202
5 (= Highest)	2.39	1.53	46
Ideology***			
No Ideology	3.26	2.11	276
Ideology	2.59	1.47	1,094
Notes: This tables provides summary statistics for inflation expectations along various dimensions. The inflation expectations were surveyed in April 2009 and pertain to the expected change in Dutch consumer prices for the next twelve months. For the income variable, the total scale (which ranges between 1 and 12) was mapped into three groups: low (values 1 to 4), medium (values 5 to 8), and high (values 9 to 12). The summary statistics may differ between this table and table 1, as not every participant in our survey returned inflation expectations. For binary dummies, *, **, and *** denote significant differences at the 10, 5, and 1 percent level.			

individuals with high levels of education expect lower increases in consumer prices ($p < 0.05$). We also find that people who have not thought about political ideologies expect higher levels of inflation ($p < 0.01$). There are also indications that expectations decline with income and social status, but these coefficients are not significant.

Next, we turn to a more formal analysis of expectations. Our survey does not enable us to explore time variation in inflation expectations. However, we have a very rich set of background variables, which we can use to explain the cross-sectional variation in expectations. We address two questions. First, does knowledge about monetary policy help in generating realistic inflation expectations? Second, does knowledge help in reducing absolute forecast errors? These two questions are related, but using the absolute errors is a stricter evaluation of the numbers submitted by the participants. At the same time, using absolute forecast errors enables us to use the variation in the data more extensively. It also enables us to assess the magnitude of the economic effect of knowledge more directly.

To define “realistic” expectations, we use several cut-off points. Looking at figure 8, some respondents clearly expect a too-high level of inflation. For instance, this would certainly apply to the nineteen individuals who expect a price increase of 10 percent over the next twelve months. What about low levels of expected inflation? For instance, 2 percent may not seem unrealistic. On the other hand, at the time of our survey, professional forecasters were expecting Dutch inflation to be moderate, with the consensus lying around 1 percent.²⁰ In the end, the actual rate of increase in Dutch consumer prices between April 2009 and April 2010 turned out to be 1.1 percent. From that perspective, an expected price increase of 2 percent may already be considered too high. As the threshold between realistic and unrealistic expectations is not clear, we use two cut-off points: 1 percent and 2 percent, meaning that unrealistic expectations start at 2 and 3 percent, respectively.²¹ Under these two definitions, the percentage of individuals with realistic expectations is 17.8 percent and 55.9 percent, respectively. Using the absolute forecast error is more straightforward. We use the actual rate of increase in Dutch consumer prices as the realized measure of inflation, so that the absolute forecast error equals $|\pi^e - 1.1|$, where π^e denotes the submitted inflation expectation.

Table 9 shows the results for these two exercises. Columns 1–4 show marginal effects for probit regressions, where the dependent variable equals 1 if the respondent has realistic expectations, and 0 otherwise. The first two columns use the threshold of 1 percent, while columns 3 and 4 use the threshold of 2 percent. Columns 5 and 6 use the absolute forecast error as the dependent variable; the results are coefficients from OLS regressions. The key explanatory variable is the level of knowledge about ECB monetary policy. The first measure we use is the correct number of answers to statements about the ECB’s objective (columns 1, 3, and 5). The second measure is the fraction of correctly answered true and false statements (columns 2, 4, and 6). As control variables, we include all the

²⁰In March 2009, the Netherlands Bureau of Economic Analysis was projecting an increase of consumer prices of 1 percent for 2009 and 2010. The April 2009 Consensus Forecasts for these two years were in the same range.

²¹We also investigated thresholds at 3 and 4 percent. These results (available on request) were similar to those with the threshold at 2 percent.

Table 9. Knowledge and Inflation Expectations

	Realistic Expectations				Absolute Forecast Error	
	Ending at 1%		Ending at 2%		(5)	(6)
	(1)	(2)	(3)	(4)		
Knowledge:						
Number of Correctly Answered Questions	0.01** (0.01)					
Fraction of Correctly Answered True Statements		-0.01 (0.04)	0.02*** (0.01)	0.02 (0.05)	-0.07*** (0.02)	-0.14 (0.16)
Fraction of Correctly Answered False Statements		0.13*** (0.04)		0.18*** (0.06)		-0.58*** (0.16)
Other Variables:						
Desire to be Informed (D)	0.04*** (0.02)	0.05*** (0.02)	0.02 (0.02)	0.03 (0.02)	-0.13** (0.07)	-0.14** (0.07)
SI1: Income	-0.00 (0.01)	-0.00 (0.01)	0.01 (0.01)	0.01 (0.01)	-0.03 (0.02)	-0.03 (0.02)
No Ideology (ID)	0.00 (0.04)	-0.00 (0.04)	-0.09** (0.05)	-0.10** (0.05)	0.69*** (0.18)	0.70*** (0.18)
Work Related to Monetary Issues	-0.00 (0.01)	-0.00 (0.01)	-0.01 (0.01)	-0.01 (0.01)	0.03 (0.03)	0.03 (0.03)
Age	-0.00 (0.00)	-0.00 (0.00)	-0.00** (0.00)	-0.00** (0.00)	0.01*** (0.00)	0.01*** (0.00)

(continued)

Table 9. (Continued)

	Realistic Expectations				Absolute Forecast Error	
	Ending at 1%		Ending at 2%		(5)	(6)
	(1)	(2)	(3)	(4)		
Status	0.02 (0.01)	0.02 (0.01)	0.03* (0.02)	0.03* (0.02)	-0.10** (0.05)	-0.10** (0.05)
City	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	0.07* (0.04)	0.07* (0.04)
TV	0.02 (0.04)	0.03 (0.04)	0.06 (0.05)	0.07 (0.05)	0.04 (0.17)	0.02 (0.17)
Newspapers	0.03 (0.05)	0.03 (0.05)	0.08 (0.05)	0.08 (0.05)	-0.15 (0.16)	-0.16 (0.16)
Other	0.03 (0.05)	0.04 (0.06)	0.04 (0.06)	0.04 (0.06)	-0.13 (0.19)	-0.14 (0.19)
Constant					1.74*** (0.33)	1.73*** (0.33)
Probability [realistic = 1]	0.18	0.18	0.58	0.58		
McKelvey and Zavoina's	0.02	0.03	0.01	0.01		
Pseudo R ²					0.06	0.06
Adjusted R ²						

Notes: Robust standard errors are shown in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Columns 1–4 present marginal effects for probit regressions. The dependent variables are binary dummies that indicate whether or not an individual's inflation expectations were realistic. Expectations are defined as unrealistic if they are above 1 and 2 percent, respectively. Marginal effects are computed at sample means, except for (0,1) dummies. The measure of media use included is the most important source of information (S). Columns 5 and 6 present coefficients from OLS regressions, where the absolute forecast error is the dependent variable. Parameters for education, sex, and partner are not significant and are not shown in this table. Results are based on 1,104 observations.

variables that were used in table 5 to explain knowledge. Finally, we include the degree to which respondents' work is related to economic, monetary, or financial issues, as this variable may be relevant for the extent to which realistic expectations are formed. The included media variables are those that measure the most important source of information (S). Results using Q , or QH and QL , were similar (available on request).

The regression results suggest that knowledge about monetary policy matters. Columns 1–4 indicate that the respondents' ability to formulate realistic inflation expectations is positively related to their level of knowledge about the ECB's objectives. Respondents who gave a higher number of correct answers also have a higher probability of giving realistic figures for inflation over the next twelve months (columns 1 and 3). Starting from the mean number of correct answers, an additional correct answer would increase the probability of realistic expectations with 1 to 2 percent. Again, the difference between true and false statements turns out to be important. As shown in the second and fourth columns of table 9, it is mainly the individual's ability to answer false statements correctly that is positively related to the ability to formulate realistic expectations. A similar picture emerges for the absolute forecast error (columns 5 and 6). Answering an additional question correctly is associated with a decrease of the forecast error by 0.07 percentage points. Again, the correlation between knowledge and expectations is the strongest for respondents who correctly answer the false statements.

Turning to the other variables, we find significant results for the desire to be informed, having no ideology, and age. The threshold for realistic inflation expectations is relevant. If the threshold for realistic expectations is set at 1 percent, the desire to be informed is significant, whereas for higher thresholds, ideology and age are significant. As 1 percent is the strictest criterion, it is perhaps not surprising that precisely those individuals who have a strong motivation to obtain information have a higher likelihood of formulating realistic expectations. For less strict definitions of realistic expectations, individuals who have not thought about ideology have about a 10 percent lower probability of formulating realistic expectations than those who have thought about this subject. Also, the elderly have a somewhat lower probability of formulating realistic expectations. We also find that age is positively correlated to the absolute forecast

error. Finally, it is remarkable that the degree to which respondents' work is related to economic, monetary, or financial issues is neither related to the realism of inflation expectations nor to the absolute forecast error. Intuitively, we would have expected that the extent to which people's jobs are related to economic issues would have been relevant for the quality of their inflation expectations.²²

Finally, we investigate whether particular pieces of information are particularly useful for the general public. To this end, we aggregated the number of correct answers to questions for five categories: price stability (statements 1, 2, and 7), economic growth (statements 3 and 4), unemployment (statements 5 and 6), countries or euro area (statements 9 and 10), and other issues (statements 8 and 11). Using these variables, we again test what determines realistic inflation expectations and absolute forecast errors. Table 10 shows the results. The useful elements of knowledge are related to economic growth and price stability. When using realistic inflation expectations (using 1 percent or 2 percent as a cut-off point) as dependent variables, the number of correct answers to questions on economic growth has a strong connection to the quality of expectations (columns 1 and 2). Moreover, correctly answering questions on economic growth and price stability is negatively related to the absolute forecast error. From the perspective of central bank communication, the issues of economic growth and price stability (and the possible trade-offs between them) would therefore be good candidates for providing greater clarity to the general public.

7. Conclusions and Policy Implications

Does the general public know what central banks do? Is this kind of knowledge relevant? Based on our analysis, we would answer these questions as follows: "partially, at most" and "yes, it is." So far, research on central bank transparency and communication has focused on the impact of communication on financial markets. We report the outcomes of a survey that examines how well the general

²²As a robustness check, we reran the regressions in table 9 without the desire to be informed, as this variable may be strongly correlated to having a job in economics. However, the coefficient for a job related to monetary issues remained insignificant (results available on request).

Table 10. Which Types of Knowledge Are Most Relevant for Inflation Expectations?

	Realistic Expectations		Absolute Forecast Error
	Ending at 1%	Ending at 2%	
	(1)	(2)	(3)
Knowledge About:			
Price Stability	0.01 (0.02)	0.02 (0.02)	−0.14** (0.06)
Unemployment	−0.01 (0.01)	−0.00 (0.02)	0.00 (0.06)
Economic Growth	0.04*** (0.02)	0.07*** (0.02)	−0.22*** (0.06)
Countries or Euro Area	0.00 (0.02)	−0.00 (0.02)	0.02 (0.07)
Other Issues	0.01 (0.02)	0.01 (0.03)	0.03 (0.07)
Other Variables:			
Desire to be Informed (D)	0.04*** (0.02)	0.03 (0.02)	−0.14** (0.06)
SI1: Income	−0.00 (0.01)	0.01 (0.01)	−0.03 (0.02)
No Ideology (ID)	0.01 (0.04)	−0.09* (0.05)	0.67*** (0.18)
Work Related to Monetary Issues	−0.01 (0.01)	−0.02 (0.01)	0.04 (0.03)
Age	−0.00 (0.00)	−0.00** (0.00)	0.01*** (0.00)
Status	0.02 (0.01)	0.03* (0.02)	−0.09* (0.05)
City	−0.01 (0.01)	−0.01 (0.01)	0.06* (0.04)
TV	0.03 (0.04)	0.07 (0.05)	0.00 (0.17)
Newspapers	0.03 (0.05)	0.09* (0.05)	−0.18 (0.16)
Other	0.04 (0.06)	0.05 (0.06)	−0.17 (0.19)
Constant			1.75*** (0.34)
Probability [realistic = 1]	0.18	0.58	
McKelvey and Zavoina’s Pseudo R ²	0.03	0.04	
Adjusted R ²			0.07

Notes: Robust standard errors are shown in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. The knowledge variables are measured as the number of correct answers per category. For the other variables, see the notes to table 9.

public understands the objectives of the ECB. We not only assess the respondents' understanding but also examine whether the channels they use to obtain information about monetary policy affect their understanding. In general, more intensive use of information results in better knowledge, suggesting that the media channel may play an important and constructive role in enhancing the public's knowledge.

There is a lot of potential to build knowledge, as we find that understanding of monetary policy is far from perfect. The average number of correct answers to eleven statements about the ECB's objectives is less than five. One particular issue is that many respondents think that the ECB's inflation target applies to *individual* countries. This finding is potentially worrisome. Even when the ECB meets its inflation target, it may face criticism in countries where national inflation rates deviate from the "close to but below two percent" target.

Our research also suggests that several survey participants have only a weak desire to be informed. Furthermore, those respondents who believe they have adequate knowledge fail to fully understand the ECB's objectives. Both a weak desire to be informed and unawareness of insufficient knowledge constitute important barriers to improve the public's understanding. Although central bank communication can be used as a tool to improve knowledge, it will be difficult to reach those members of the public who see no clear reasons to listen. Therefore, to overcome the various barriers, central bank communication should not only focus on content in order to improve knowledge. Rather, it should also aim to convince people of the importance of monetary policy (see also Ehrmann and Tzamourani 2012).

To what extent is knowledge about monetary policy relevant for the general public? First, as noted by Blinder et al. (2008), it is the public that gives central banks their democratic legitimacy, and hence their independence. To this end, it is crucial that the public understands what monetary authorities do. Our paper presents a second dimension to the case for adequate understanding, as we find evidence that the level of knowledge on monetary policy has economic implications. Our results indicate that individuals with better knowledge about monetary policy objectives have a higher

probability of formulating realistic inflation expectations. Also, we document a negative relationship between monetary policy knowledge and inflation forecast errors. Knowledge on price stability and economic growth is particularly strongly related to the quality of inflation expectations. These findings suggest there are important returns to knowledge on monetary policy, as the ability to make a realistic assessment of future price developments is crucial for adequate household financial decision making. It would be interesting to explore the effects of knowledge about monetary policy along other economic dimensions. We leave this for future research.

Appendix 1. The Questionnaire

Intro 1: This questionnaire concerns the European Central Bank, but we will first ask you three general questions.

Q1. Has the current economic and financial crisis had negative consequences for you, or do you expect such consequences in the near future?

Yes, very much
Yes, somewhat
Hardly
No
I don't know

Q2. Does (or did) your work relate to economic, monetary or financial issues?

Yes, almost every day
Often, but not on a daily basis
Sometimes
Hardly
Never

Q3. How would you describe your political orientation? (you may choose more than one option)

Liberal
 Socialist
 Christian democrat
 Conservative
 Progressive
 I haven't thought about it
 Other, namely ...

Intro 2: The remainder of this questionnaire is on the European Central Bank (ECB). The ECB is the central bank for Europe's single currency, the euro. Since 1999 the euro has been introduced in 16 European countries.

Q4. How do you judge your own knowledge about...

	Very poor	Poor	Neutral	Good	Very good	I don't know
... the European Central Bank						
... economic developments (like price changes, economic growth and unemployment)						

Q5. How important is it to you to be well informed on the policies of the European Central Bank (ECB)?

Not important at all
 Not very important
 Somewhat important
 Very important
 Extremely important
 I don't know

[Note: respondents who answered "not important at all" continued with Q7.]

Q6. Could you please tell us more about your reasons to be informed about the policy of the European Central Bank (ECB)?

	Not important at all	Not very important	Somewhat important	Very important	Extremely important	I don't know
ECB policy affects my personal or family income.						
ECB policy influences how much my money can buy.						
ECB policy affects my business/job/profession.						
ECB policy affects the value of my stocks or other investments.						
ECB policy is important for the economy.						
I just like to keep informed well.						
Other, namely:						

Intro 3: We would like to gain insight into people's understanding of the most important goal of the European Central Bank. There is no need to search for the correct answers, and you should not worry about giving an incorrect answer.

Q7. Can you please indicate whether the following statements that refer to the main objective of the European Central Bank (ECB) are true or false?

The main objective of the ECB...	True	False	I don't know
1... is price stability			
2... is to keep prices constant			
3... is low unemployment			
4... is an unemployment rate of at most 5%			
5... is high economic growth			
6... is an economic growth rate of at least 2%			

- 7... is an inflation rate that is close to but below 2%
- 8... is to keep interest rates constant
- 9... applies to the euro area average
- 10... applies to all euro area countries separately
- 11... applies to the medium-term

Intro 4: We would like to ask you a number of questions concerning the manner in which you receive information about the policy of the European Central Bank (ECB).

Q8. Through which sources of information do you hear or read about the policies of the European Central Bank (ECB)?

Regularly Occasionally Never I don't know

Television

Radio

Newspapers

Magazines

Internet

Friends, family, colleagues

Other, namely:

Intro 5: Finally, we would like to ask you a number of questions on the sources of information on the policy of the European Central Bank (ECB).

Q9. Which of the following is your most important source of information on the policy of the European Central Bank? (please select only one)?

Television

Radio

Newspapers

Magazines

Internet

Friends/relatives/colleagues

I don't use any sources on ECB policy

I don't know

Other, namely:

Appendix 2. Weighting

Table 1 indicates issues regarding representativeness. We have used information from Statistics Netherlands on gender, age, household size, partner, and education to compute sampling weights. The differences between the weighted and unweighted levels of knowledge are small. To be precise, when we weight observations, we find that the estimated level of knowledge is slightly lower. For instance, the average number of correct answers to question 7 of our survey declines from 4.60 to 4.45. The overall impression that the public's knowledge on monetary policy is far from perfect is left undisturbed by weighting. Unfortunately, we are unable to weight according to income. As income turns out to have a positive relationship with knowledge, we would rate the conclusions regarding the level of knowledge as being rather too optimistic than pessimistic.

There has been an extended discussion on weighting in regression analysis. For various positions in this debate we refer to DuMouchel and Duncan (1983), Pfeffermann (1993), Winship and Radbill (1994), Korn and Graubard (1995), Magee, Robb, and Burbidge (1998), and Angrist and Pischke (2009). The consensus seems that weighting is more relevant for estimating population parameters than for establishing measures of association between variables. Also, weighting may have the drawback of increasing the standard errors of the estimates.

For a number of reasons, we base our regressions on unweighted data. First, we note that the descriptive results are already not very sensitive to weighting. Second, our sampling weights are solely based on independent variables already included in the regressions. In this case, unweighted and weighted regressions will yield consistent estimates, but the former will be, as noted, more efficient (Winship and Radbill 1994). Third, we followed DuMouchel and Duncan's (1983) suggestion to include sampling weights and interaction terms between weights and independent variables in the regressions to detect possible misspecification. In almost all cases, we were not able to reject the hypothesis that the coefficients for the sampling weights and the interactions terms equal zero, which indicates using unweighted data is appropriate. Further details regarding weighting are available on request.

References

- Angrist, J., and J-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Berger, H., M. Ehrmann, and M. Fratzscher. 2006. "Monetary Policy in the Media." ECB Working Paper No. 679.
- Blanchflower, D. G., and R. Kelly. 2008. "Macroeconomic Literacy, Numeracy and the Implications for Monetary Policy." Available at <http://www.bankofengland.co.uk/archive/Documents/historicpubs/speeches/2008/speech346paper.pdf>.
- Blinder, A. S., M. Ehrmann, M. Fratzscher, J. de Haan, and D. Jansen. 2008. "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence." *Journal of Economic Literature* 46 (4): 910–45.
- Blinder, A. S., and A. B. Krueger. 2004. "What Does the Public Know About Economic Policy, and How Does It Know It?" *Brookings Papers on Economic Activity* 1: 327–87.
- Branch, W. A. 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.
- Brand, C., D. Buncic, and J. Turunen. 2010. "The Impact of ECB Monetary Policy Decisions and Communication on the Yield Curve." *Journal of the European Economic Association* 8 (6): 1266–98.
- Bruine de Bruin, W., W. Vanderklaauw, J. S. Downs, B. Fischhoff, G. Topa, and O. Armantier. 2010. "Expectations of Inflation: The Role of Demographic Variables, Expectation Formation, and Financial Literacy." *Journal of Consumer Affairs* 44 (2): 381–402.
- Bryan, M. F., and G. Venkatu. 2001. "The Demographics of Inflation Opinion Surveys." *Economic Commentary* (Federal Reserve Bank of Cleveland) (October).
- Buliř, A., M. Čihák, and D. Jansen. 2013. "What Drives Clarity of Central Bank Communication About Inflation?" *Open Economies Review* 24 (1): 125–45.

- . 2014. “Does the Clarity of Inflation Reports Affect Volatility in Financial Markets?” IMF Working Paper No. 14/175.
- Cameron, A. C., and P. K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge, UK: Cambridge University Press.
- . 2007. *Regression Analysis of Count Data*. Cambridge, UK: Cambridge University Press.
- Capistrán, C., and A. Timmermann. 2009. “Disagreement and Biases in Inflation Expectations.” *Journal of Money, Credit and Banking* 41 (2–3): 365–96.
- Carroll, C. D. 2003. “Macroeconomic Expectations of Households and Professional Forecasters.” *Quarterly Journal of Economics* 118 (1): 269–98.
- Dincer, N., and B. Eichengreen. 2007. “Central Bank Transparency: Where, Why and with What Effects?” NBER Working Paper No. 13003.
- . 2009. “Central Bank Transparency: Causes, Consequences and Updates.” NBER Working Paper No. 14791.
- Dovern, J., U. Fritsche, and J. Slaálek. 2012. “Disagreement Among Forecasters in G7 Countries.” *Review of Economics and Statistics* 94 (4): 1081–96.
- DuMouchel, W. H., and G. J. Duncan. 1983. “Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples.” *Journal of the American Statistical Association* 78 (383): 535–43.
- 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.
- Ehrmann, M., and P. Tzamourani. 2012. “Memories of High Inflation.” *European Journal of Political Economy* 28 (2): 174–91.
- European Commission. 2010. “Europeans’ Knowledge of Economic Indicators.” Survey. Available at http://ec.europa.eu/public_opinion/archives/ebs/ebs.323_en.pdf.
- Geraats, P. M. 2009. “Trends in Monetary Policy Transparency.” *International Finance* 12 (2): 235–68.
- Grabe, M. E., R. Kamhawi, and N. Yegiyani. 2009. “Informing Citizens: How People with Different Levels of Education Process Television, Newspapers, and Web News.” *Journal of Broadcasting and Electronic Media* 53 (1): 90–111.

- Greene, W. H. 2003. *Econometric Analysis*. 5th edition. Upper Saddle River, NJ: Prentice Hall.
- Hayo, B., and E. Neuenkirch. 2014. "The German Public and Its Trust in the ECB: The Role of Knowledge and Information Search." *Journal of International Money and Finance* 47: 286–303.
- Hurd, M., M. C. J. van Rooij, and J. Winter. 2011. "Stock Market Expectations of Dutch Households." *Journal of Applied Econometrics* 26 (3): 416–36.
- International Monetary Fund. 2000. *Supporting Document to the Code of Good Practices on Transparency in Monetary and Financial Policies*. Washington, DC: International Monetary Fund.
- Jansen, D. 2011. "Does the Clarity of Central Bank Communication Affect Financial Markets? Evidence from Humphrey-Hawkins Testimonies." *Contemporary Economic Policy* 29 (4): 494–509.
- Jansen, D., and J. de Haan. 2005. "Talking Heads: The Effects of ECB Statements on the Euro–Dollar Exchange Rate." *Journal of International Money and Finance* 24 (2): 343–61.
- Jansen, D., R. H. J. Mosch, and C. A. B. van der Cruysen. 2015. "When Does the General Public Lose Trust in Banks?" Forthcoming in *Journal of Financial Services Research*.
- Jonung, L. 1981. "Perceived and Expected Rates of Inflation in Sweden." *American Economic Review* 71 (5): 961–68.
- Juster, F. T. 1981. "An Expectational View of Consumer Spending Prospects." *Journal of Economic Psychology* 1 (2): 87–103.
- Katona, G. 1975. *Psychological Economics*. New York: Elsevier.
- Korn, E. L., and B. I. Graubard. 1995. "Examples of Differing Weighted and Unweighted Estimates from a Sample Survey." *The American Statistician* 49 (3): 291–95.
- Lambert, D. 1992. "Zero-Inflated Poisson Regression with an Application to Defects in Manufacturing." *Technometrics* 34 (1): 1–14.
- Lusardi, A., and O. S. Mitchell. 2007a. "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth." *Journal of Monetary Economics* 54 (1): 205–24.
- . 2007b. "Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education." *Business Economics* 42 (1): 35–44.

- Lusardi, A., and P. Tufano 2009. "Debt Literacy, Financial Experiences, and Overindebtedness." NBER Working Paper No. 14808.
- Magee, L., A. L. Robb, and J. B. Burbidge. 1998. "On the Use of Sampling Weights When Estimating Regression Models with Survey Data." *Journal of Econometrics* 84 (2): 251–71.
- Mankiw, N. 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.
- Marzilli Ericson, K. M. 2011. "Forgetting We Forget: Overconfidence and Memory." *Journal of the European Economic Association* 9 (1): 43–60.
- Mullahy, J. 1986. "Specification and Testing of Some Modified Count Data Models." *Journal of Econometrics* 33 (3): 341–65.
- Neuenkirch, M. 2012. "Managing Financial Market Expectations: The Role of Central Bank Transparency and Central Bank Communication." *European Journal of Political Economy* 28 (1): 1–13.
- Orphanides, A., and J. Williams 2004. "Imperfect Knowledge, Inflation Expectations, and Monetary Policy." In *The Inflation-Targeting Debate*, eds. B. S. Bernanke and M. Woodford. Chicago: NBER/University of Chicago Press.
- . 2005. "Inflation Scares and Forecast-Based Monetary Policy." *Review of Economic Dynamics* 8 (2): 498–527.
- Papke, L. E., and J. M. Woolridge. 1996. "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates." *Journal of Applied Econometrics* 11 (6): 619–32.
- Pfeffermann, D. 1993. "The Role of Sampling Weights When Modeling Survey Data." *International Statistical Review* 61 (2): 317–37.
- Sibert, A. 2009. "Is Transparency About Central Bank Plans Desirable?" *Journal of the European Economic Association* 7 (4): 831–57.
- Souleles, N. S. 2004. "Expectations, Heterogeneous Forecast Errors and Consumption: Micro Evidence from the Michigan Consumer Sentiment Surveys." *Journal of Money, Credit and Banking* 36 (1): 39–72.
- Teppa, F., and C. Vis. 2012. "The CentERpanel and the DNB Household Survey: Methodological Aspects." *DNB Occasional Study* 10 (4).

- van der Cruijssen, C. A. B., and M. Demertzis. 2011. "How Are Inflation Expectations Anchored in EMU Countries?" *Economic Modelling* 28 (1–2): 281–98.
- van der Cruijssen, C. A. B., and S. C. W. Eijffinger. 2010a. "From Actual to Perceived Transparency: The Case of the European Central Bank." *Journal of Economic Psychology* 31 (3): 388–99.
- . 2010b. "The Economic Impact of Central Bank Transparency: A Survey." In *Challenges in Central Banking: The Current Institutional Environment and Forces Affecting Monetary Policy*, ed. P. Siklos, M. Bohl, and M. Wohar. New York: Cambridge University Press.
- van der Cruijssen, C. A. B., D. Jansen, and J. de Haan. 2010. "How Much Does the Public Know About the ECB's Monetary Policy? Evidence from a Survey of Dutch Households." DNB Working Paper No. 252.
- van Rooij, M. C. J., A. Lusardi, and R. J. M. Alessie. 2011. "Financial Literacy and Stock Market Participation." *Journal of Financial Economics* 101 (2): 449–72.
- . 2012. "Financial Literacy, Retirement Planning and Household Wealth." *Economic Journal* 122 (560): 449–78.
- Wärneryd, K. E., and R. Wahlund. 1985. "Inflationary Expectations." In *Economic Psychology*, ed. H. Brandstätter and E. Kirchler. Linz: Trauner.
- Winkelmann, R. 2008. *Econometric Analysis of Count Data*. 5th edition. Berlin: Springer Verlag.
- Winship, C., and L. Radbill. 1994. "Sampling Weights and Regression Analysis." *Sociological Methods and Research* 23 (2): 230–57.

Model-Based Ex Post Evaluation of Monetary Policy*

Eyal Argov, Alon Binyamini, Eliezer Borenstein, and
Irit Rozenshtrom

Research Department, Bank of Israel

We present a model-based methodology to conduct an ex post evaluation of monetary policy decisions, by testing whether alternative policy decisions could have brought a Pareto improvement in terms of inflation and output volatilities. This involves simulations of counterfactual scenarios under alternative monetary policy shocks, and computation—for each such simulation—of the root mean square (RMS) of the inflation and output gaps during and following the evaluated year. It is then possible to compare the actual RMS with simulation-based frontiers, with each frontier reflecting different constraints on interest rate volatility, which can be viewed as a third objective variable. The actual RMS is also compared with the counterfactual RMS that would have been obtained under the case of no policy shocks. Such comparisons enable testing whether monetary policy shocks were “ex post efficient.” The methodology is implemented in an evaluation of Bank of Israel policy decisions during the years 2001–11. The implementation shows several distinct sets of years: years in which actual RMSs were close to the efficient frontiers and years in which they were distant from them; years in which monetary policy shocks led to an absolute improvement in economic outcomes (by reducing the RMSs of both inflation and output gaps) or an absolute deterioration; and years in which policymakers faced a trade-off between all three objective variables or

*The authors would like to thank Guy Segal, Roni Frish, participants at the Bank of Israel Research Department seminar, two anonymous referees, and the co-editor for their valuable comments. Author e-mails: eyal.argov@boi.org.il, alon.binyamini@boi.org.il, eliezer.borenstein@boi.org.il, and irit.rozenshtrom@boi.org.il.

between a subset of the variables. For most evaluated years, the results seem qualitatively robust, considering the uncertainty in the historical shocks extracted, as well as in alternative definitions for the output gap.

JEL Codes: C54, E37, E52, E58.

1. Introduction

The purpose of this paper is to offer an analytical framework for an ex post evaluation of the efficiency of monetary policy decisions. A motivation for such a framework is discussed by Svensson (2012), who argues that accountability and regular evaluation of the central bank's decisions are important incentives in motivating independent and powerful central banks. To that end, he encourages the development and application of reasonable principles and appropriate methods.

To the best of our knowledge, there has been little work done on principles and methods for such an ex post evaluation of monetary policy decisions. The vast literature on optimal policy *rules* focuses on the general functioning of decision makers, i.e., evaluating the policy rule, rather than the actual decisions.¹ One exception is Svensson (2012), who proposes a framework for evaluating a specific policy decision. This involves calculating the expected root mean squares (RMSs) of the inflation and output gaps ex ante, as they were projected when the interest rate decisions were made. Expected RMSs are also calculated for alternative future interest rate paths using a DSGE model and then compared with the baseline case. Such a procedure allows the quantifying of the trade-off between stabilizing inflation and output, and may show at times that the expected RMSs of both inflation and output could have been reduced, and therefore that the policy decisions were ex ante inefficient.

The approach taken here differs from that of Svensson (2012) in two main ways: First, the approach can be applied in the evaluation of a sequence of policy decisions, say over a period of a year, and not just of a projected path associated with a single decision.

¹Using the terminology of DSGE models, the latter includes not only the systematic part of the rule but also policy “shocks,” i.e., non-systematic deviations from the rule.

Second, and more importantly, we employ an approach that evaluates past monetary policy decisions based on ex post information, which was not always available in real time. We nevertheless believe that it is of interest to examine what would have been the counterfactual outcomes under alternative decisions, what would have been the best economic results that policymakers could have delivered, and what are the measures that should have been taken to achieve those results. Moreover, a single model clearly cannot forecast certain developments that may be predictable using other tools. Hence, the model employed may regard such developments as unexpected shocks (revealed only ex post), whereas in reality they could have been predicted and taken into account by decision makers. Put differently, to some extent, ex post evaluation may point to conclusions relevant ex ante as well.

The proposed evaluation framework compares the actual volatility of two main objective variables—the inflation and output gaps—to their counterfactual volatility associated with alternative interest rate paths. The analysis uses a DSGE model, and the volatility is measured by the RMS of a variable during the evaluated period and its expected path for subsequent periods. It is the effect of policy decisions on future outcomes that motivates the inclusion of the expected path in the calculation of the RMS. The counterfactual interest rate paths are obtained by repeatedly hitting the model economy with randomly selected monetary policy shocks. One particular benchmark alternative path is the case of zero policy shocks, i.e., strictly following the interest rate rule. In addition, we numerically search for the efficient frontier, i.e., the combinations for which the RMSs of both objective variables cannot be simultaneously improved. We regard the “active component” of the policy (that is, the non-systematic policy shock) as “efficient” when actual RMSs are closer to the frontier, compared with those consistent with the zero-policy-shock path. In addition, the evaluation framework can be used to assess whether the actual policy was more in favor of stabilizing inflation or stabilizing activity.

The main underlying assumption of this evaluation framework is that, over short horizons, an active policy means choosing deviations from an interest rate rule, i.e., manipulating the monetary policy shocks, rather than changing the policy rule per se. This assumption stems from our view that the systematic component

of monetary policy rules cannot, and should not, be changed frequently. One reason is that central banks do not normally publish their policy rule—certainly not on a quarterly basis. If anything, central banks publish an expected policy *path*. Therefore, considering counterfactual outcomes for alternative policy *rules* may not be particularly relevant. Since frequently changing the policy rule is not practical, and since the public presumably employs some form of assumed rule to form expectations with regard to future economic developments, we interpret active monetary policy to be the choice of short-run deviations from the rule, i.e., monetary policy “shocks.” The driving forces behind such shocks may be (temporary) changes in the loss function’s weights (such as a change in the relative importance assigned to stabilizing inflation or output, as well as the weight assigned to interest rate smoothness) or considerations that are not systematically incorporated into the rule (for example, precautionary reaction to risk). It should be noted that there are, of course, alternative interpretations for the deviations from an interest rate rule, which are not consistent with our methodology; these include real-time measurement errors in the variables entering the rule and errors resulting from misspecification of the “policy rule.”

Our approach thus focuses on the discretionary, or the non-systematic, part of the monetary policy. Such an approach is consistent with the view expressed by Bernanke and Mishkin (1992), who stress that rules do not always allow for responding to unforeseen circumstances. It is also consistent with Svensson (2005), who argues that monetary policy which incorporates judgment may perform better than simply following an instrument rule. It is exactly this view of the non-systematic, or discretionary, component that makes our evaluation framework relevant.

In contrast to our focus on the non-systematic component, Taylor (2011a, 2011b) emphasizes the important stabilizing role of rule-based policy, that is, of the systematic part of the policy reaction function. This view stresses the expectations channels of the monetary policy transmission mechanism.² In this context, it is

²Taylor (2011b) argues that the first decade of the 2000s was characterized by great deviations from the rule-based policy, a rule that was successful in stabilizing economic activity, thus attaining the so-called Great Moderation. He further argues that it was this great deviation which was responsible for the Great Recession that followed. Note that this view fits well into our approach.

worth noting that we assume that the policy shocks, which reflect the discretionary component of the policy, do not impair the credibility of the policymaker. Thus, the rule-based part of the policy continues to have an effective stabilizing role.

In the proposed framework, the criterion for policy evaluation focuses on inflation- and output-gap volatilities. However, it is common to suppose that monetary policy attributes some value to interest rate smoothness as well. Therefore, we compute different efficient frontiers, corresponding to different limitations on interest rate volatility. This allows the analysis to introduce, in some sense, a third argument into the implicit loss function.

To demonstrate an implementation of the framework, we employ the Bank of Israel's medium-scale DSGE model (MOISE)³ to evaluate the conduct of monetary policy in Israel during the years 2001–11. During this decade, inflation in Israel fluctuated around the center of the target range, though there were sizable deviations above and below the midpoint. For most of the period, the (model-based) output gap was negative (2 percent on average), which in retrospect can be viewed as an indication that there was considerable room for further monetary easing, namely policy shocks, that may have improved economic outcomes. The implementation reveals several distinct sets of years: years in which actual RMSs were close to the efficient frontier and years in which they were distant; years in which monetary policy shocks led to an absolute improvement in economic outcomes or to an absolute deterioration; and years characterized by aggressive policy shocks, which were usually aimed at narrowing the output gap at the expense of more volatile inflation. The implementation is followed by robustness checks in which we consider the sensitivity of the conclusions to the extraction of the historical shocks and to alternative definitions for the output gap. For most evaluated years, these robustness checks suggest that the results remain qualitatively unchanged and, accordingly, the conclusions remain valid.

The rest of the paper is organized as follows. Section 2 formally describes the framework for policy evaluation. Section 3 demonstrates an implementation of the framework, by evaluating monetary

³See Argov et al. (2012).

policy in Israel during the decade from 2001 to 2011. In that section we will only elaborate on a few distinct years. In section 4 we present two robustness analyses—with respect to alternative extractions of historical shocks and with respect to alternative definitions of the output gap. Finally, section 5 offers some concluding remarks—both on the framework and on the evaluation of the Bank of Israel’s monetary policy during the first decade of the millennium as a whole.

2. The Evaluation Framework

2.1 *A Model*

This section describes the formal framework. In order to apply the model-based evaluation framework, one has to choose a model to be employed both for the extraction of historical shocks that are consistent with the observed history and for the simulation of counterfactual scenarios. Monetary policy must be explicitly specified in this model, using a rule that includes an exogenous shock. These restrictions are very general so that any commonly used DSGE model is adequate.

For the demonstration of the framework in section 3 below, we use the Bank of Israel’s medium-scale DSGE model, MOISE, described in Argov et al. (2012). The model follows along the lines of the European Central Bank’s New Area-Wide Model and the Riksbank’s Ramses model (see Christoffel, Coenen, and Warne 2008 and Adolfson et al. 2007, respectively).⁴ The economic entities in the model are households, a production sector (producers and importers of intermediate goods, producers of final goods, and exporters), a fiscal authority, and an inflation-targeting central bank whose policy tool is the nominal interest rate ($\hat{r}_t$). The interest rate rule in MOISE takes the following form (a hat over a variable denotes logarithmic deviation from a steady-state value or from a trend):

$$\begin{aligned} \hat{r}_t = & (1 - \phi_R) \left[\hat{r}_t^{fwd} + \hat{\pi}_t + \phi_\Pi \left(\hat{\pi}_t^{CB} - \hat{\pi}_t \right) + \phi_y \hat{y}_t^{GAP} + \phi_{\Delta S} \Delta S_t \right] \\ & + \phi_R \hat{r}_{t-1} + \eta_t^R, \end{aligned} \quad (1)$$

⁴Other similar central bank models include Brubakk et al. (2006) for Norway, Murchison and Rennison (2006) for Canada, Beneš et al. (2009) for New Zealand, and Seneca (2010) for Iceland.

where the inflation to which the rule responds is

$$\hat{\pi}_t^{CB} = E_t [\hat{\pi}_{C,t-2} + \hat{\pi}_{C,t-1} + \hat{\pi}_{C,t} + \hat{\pi}_{C,t+1}], \quad (2)$$

with $\hat{\pi}_{C,t}$ being the quarterly inflation of the consumer price index (CPI) in period t .

Thus, the interest rate is driven by deviations of (expected) inflation from the inflation target ($\hat{\pi}_t^{CB} - \hat{\pi}_t$), the output gap ($\hat{y}_t^{GAP}$)—defined below as the deviation of output from a technology-driven trend—and nominal depreciation (ΔS_t). The variable $\hat{r}_t^{fwd}$ is the long-run real interest rate, and $\hat{\pi}_t$ is the time-varying inflation target. The equation includes a policy shock, η_t^R , which is assumed to follow a white-noise process.

In this paper we define the output gap as the logarithmic deviation of the observed output (Y_t) from a technology-driven trend:

$$\hat{y}_t^{GAP} \equiv \log \frac{Y_t}{Z_t \varepsilon_t} - \log y, \quad (3)$$

where Z_t and ε_t are, respectively, permanent and transitory technology shocks (external shifts in the total factor productivity) and y is the steady state of the productivity-adjusted output. This definition of the output gap, which will also be used as an objective variable in the policy evaluation, is a measure of resource utilization. In MOISE it essentially accounts for deviations of production inputs (capital and labor) from some (unobserved) trend. Hence, our output-gap measure may be interpreted as reflecting demand pressures and rigidities such as “time to build,” rather than an alternative measure that uses only the permanent component of technology as the trend. There are some alternative measures for the output gap that are used in the New Keynesian literature, particularly in the literature on optimal monetary policy. These include the gap from the hypothetical flexible-price output or from perfectly competitive output. Apart from the complexities involved in calculating these theoretical measures, we suspect that using them in practice for policy evaluation might make the results more dependent on model assumptions. Svensson (2012), who uses output deviations from HP trend, stresses the importance of having several points of view with

respect to resource utilization. In line with this view, subsection 4.2 presents robustness analysis with respect to alternative measures for the output gap.

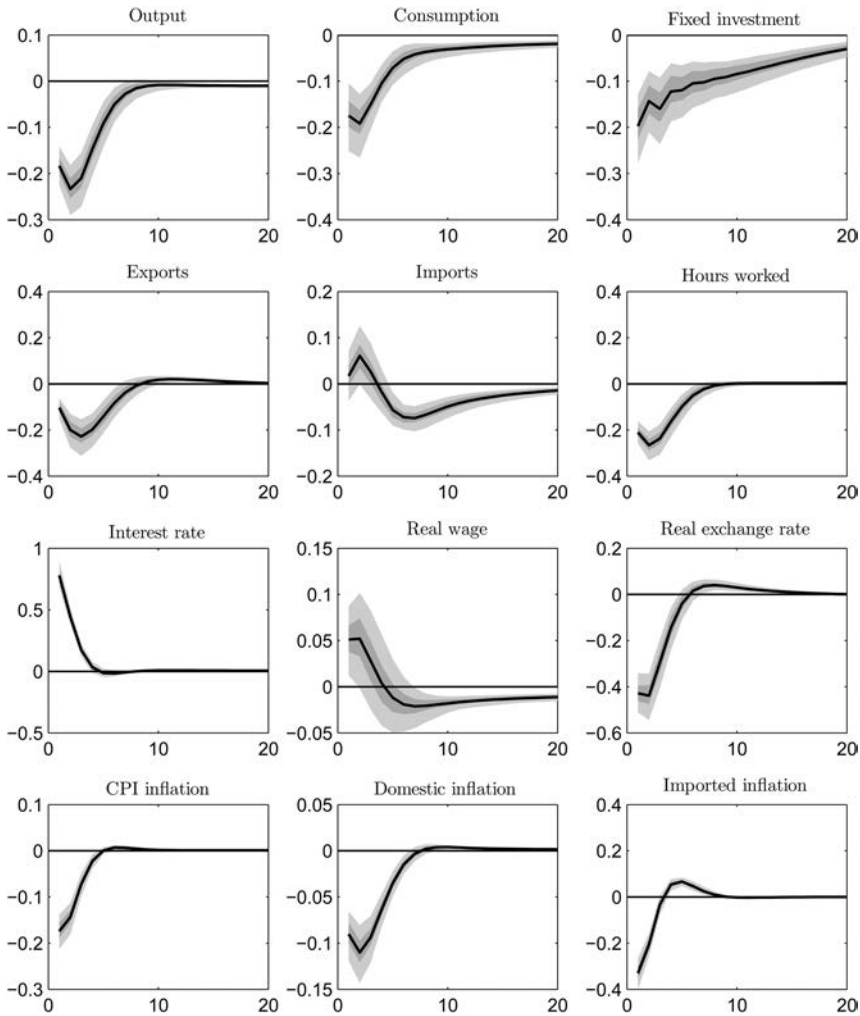
2.2 *The Monetary Policy Shock*

Since the non-systematic part of the policy rule, namely the monetary policy shock, is at the center of our evaluation approach, in this subsection we briefly discuss the effect of a monetary policy shock in the model. This shock will be manipulated within our evaluation framework, in order to simulate alternative, counterfactual outcomes. Figure 1 presents the impulse response function (IRF) for several key variables following a monetary policy shock, η^R .

As can be seen from figure 1, an innovation to the interest rate rule (1) of one standard deviation triggers an immediate increase of the interest rate of 0.75 percentage points. This increase causes output to fall by approximately 0.2 percent. Note that output reaches its lowest point only after two quarters and gradually converges back to its trend within two years. CPI inflation falls on impact by 0.2 percentage points, with an accumulated effect of about 0.4 percentage points, one year after the shock.

It is interesting to compare the model's IRF to those reported for similar models of other economies, such as in Christoffel, Coenen, and Warne (2008) and Adolfson et al. (2007) for the euro area, Adolfson et al. (2008) for Sweden, and Beneš et al. (2009) for New Zealand. This comparison leads to four general observations: (i) the size of the shock in MOISE is typically larger by a scale of 1.5 to 3 (reflecting a larger estimated standard deviation of the interest rate shock); (ii) while the effect on output in MOISE is typically smaller, mainly due to the lower sensitivity of investment, the effect on inflation is larger due to a faster exchange rate pass-through (along with higher import intensity); (iii) the output reaction in MOISE is faster and less hump shaped, with the strongest effect only two quarters following the shock, as compared with three or four quarters in other economies; and finally, (iv) the effect of the shock has a shorter duration in MOISE (two years as compared with five years in other economies).

Figure 1. Impulse Response to an Interest Rate Shock



Notes: Shock of one standard deviation. Solid line: mean of impulse response. Gray area: 70 and 90 percent highest interval of impulse response. Real variables: percentage deviation from steady state. Inflation: percentage-point deviation from steady state. Interest rate: annualized percentage-point deviation from steady state.

2.3 Formal Description of the Algorithm

This section lays out a step-by-step description of the evaluation framework. The outline of the algorithm is followed by a discussion of the typical output of the evaluation exercise.

Step 1: Estimate the Historical Shocks. The Kalman filter is used to extract historical shocks η_t for all periods t in the sample, given a set of observed variables.⁵ η_t is a vector of all the model's i.i.d. shocks including the monetary policy shock, η_t^R . Note that since the historical shocks are not uniquely identified (there are infinitely many combinations of shocks consistent with the observable variables), we employ the Kalman filter to smooth out the conditional mean of the distribution of shocks' estimates. In section 4.1 we present a robustness test, accounting for the uncertainty associated with the actual realization of these historical shocks.

Step 2: Choose the Period of Evaluation and Calculate the Actual RMSs. The period of evaluation does not necessarily have to be a year. Nonetheless, for our implementation below we choose to work with calendar years, for several reasons: (i) it is a natural reference period; (ii) it is short enough so that monetary policy may be thought of as choosing deviations from some rule and not as choosing different rules; and (iii) it is long enough for policy to have an economically significant effect on the target variables, as is evident from the impulse response presented in figure 1.⁶

For the chosen period, calculate the RMSs of the variables assumed to be targeted by monetary policy. In our implementation, we assume the policy criterion to be a loss function mainly based on the RMS of the deviations of the inflation rate from its target (inflation gap, $\hat{\pi}_t^{GAP} = \hat{\pi}_{C,t} - \hat{\pi}_t$) and deviations of output from a trend consistent with technological capacity (i.e., the output gap, $\hat{y}_t^{GAP}$). Formally, we let t be the first quarter of the evaluated year, and calculate

⁵Throughout the paper, we actually use the smoothing process—that is, estimating unobservable shocks based on the entire sample.

⁶Israel has a relatively rapid transmission mechanism. Therefore, in applying this framework to other economies, one might choose to evaluate periods of two to three years.

$$\begin{aligned}
& RMS(\hat{\pi}_{t,t+7}^{GAP}) \\
&= \sqrt{\frac{1}{8} \left[(\hat{\pi}_{C,t} - \bar{\pi}_t)^2 + \cdots + (\hat{\pi}_{C,t+3} - \bar{\pi}_{t+3})^2 \right.} \\
&\quad \left. + E(\hat{\pi}_{C,t+4} - \bar{\pi}_{t+4})^2 + \cdots + E(\hat{\pi}_{C,t+7} - \bar{\pi}_{t+7})^2 \right] } \quad (4)
\end{aligned}$$

and

$$\begin{aligned}
& RMS(\hat{y}_{t,t+7}^{GAP}) \\
&= \sqrt{\frac{1}{8} \left[(\hat{y}_t^{GAP})^2 + \cdots + (\hat{y}_{t+3}^{GAP})^2 + (E\hat{y}_{t+4}^{GAP})^2 + \cdots + (E\hat{y}_{t+7}^{GAP})^2 \right]}. \quad (5)
\end{aligned}$$

In equations (4)–(5), E is the mathematical expectation operator, conditional on the information available at the last quarter of the evaluated year, $t + 3$, so that the RMS includes a forecast for the respective variable in the following year. One reason for taking the expectations rather than future realizations (ex post) in calculating the “actual” RMS is to avoid including the effects of future policy decisions in the analysis, thus focusing on policy decisions in the period under evaluation only. Also, the further ahead into the future the realization of shocks is, the less we would expect policy-makers to anticipate them and act accordingly. Finally, this way the evaluation of policy for a historical period is similar to an evaluation one would have performed in the end of that period, bringing us closer to an evaluation performed in real time.⁷

At this stage we also calculate the within-year standard deviation of the interest rate:

$$\sigma_{t-1,t+3}^r = \sqrt{\frac{1}{5} \left[(r_{t-1} - \bar{r})^2 + (r_t - \bar{r})^2 + \cdots + (r_{t+3} - \bar{r})^2 \right]}, \quad (6)$$

where r_t denotes the annualized nominal interest rate and $\bar{r}$ is the within-year average.⁸

⁷Two reasons why the results might be different, compared with those based on real-time evaluation, are data revisions and the Kalman smoother which exploits information inherent in the entire sample, available only to the ex post researcher. We remark on the issue of data vintages further below.

⁸In fact, starting in $t - 1$, expression (6) also includes the last quarter of the preceding year, capturing the idea that the evaluated year is part of a

Step 3: Run a Counterfactual Simulation with Zero Monetary Policy Shocks. Starting from the smoothed state in the last quarter before the evaluated period, a counterfactual scenario is simulated based on the vectors of extracted shocks from step 1, except for the monetary policy shocks which are set to zero in every quarter of the evaluated period. All shocks for the periods following the evaluated one are set to zero as well. The result suggests what would have been the case had monetary policymakers not deviated from the estimated interest rate rule.

Using the simulation results, the RMSs of the main objective variables are calculated, as in equations (4) and (5). Comparing the gaps between the resulting counterfactual RMSs with the actual ones, from step 2, reflects the contribution of the policy's non-systematic component.

Step 4: Run Numerous Counterfactual Simulations with Randomly Selected Monetary Policy Shocks, and Estimate Policy Frontiers. Run numerous simulations based on counterfactual, alternative monetary policy shocks. In our implementation, 5,000 four-period vectors of monetary policy shocks are drawn from the estimated distribution of the monetary policy shock, $N\left(0, \sigma_{\eta_R}^2\right)$. For each draw, a counterfactual simulation is run with the vector of drawn monetary policy shocks and the remaining smoothed shocks extracted by step 1. For each simulation, the RMSs are calculated for the objective variables, as in equations (4) and (5).

In addition, we estimate the feasible policy frontier—the combinations of RMSs for which it is impossible to improve the RMSs of both the output gap and the inflation gap. Moreover, in order to treat interest rate volatility as an additional consideration of policymakers, we distinguish between different frontiers, each corresponding to a different ceiling on the within-year interest rate standard deviation, as calculated in equation (6). Note that the frontiers are estimated using a search algorithm (presented in the appendix), such that each frontier is based on additional simulations

continuum. That is, when deciding on the interest rate for the first quarter of the calendar year, the central bank is also constrained by the interest rate of the previous quarter. For the same reason, the “within-year average” also includes that last quarter of the preceding year.

in addition to the 5,000 random draws from the estimated distribution.⁹

Step 5: Generate a Diagrammatic Representation of the Results. In our implementation below, we present the evaluation results using two types of diagrams: (i) a scatter diagram that presents the counterfactual RMSs of the objective variables, along with feasible frontiers; and (ii) the interest rate paths of the simulations associated with one of the frontiers.

The first type of diagram (see, for example, panel A in figure 4) presents the RMSs of the inflation gap (horizontal axis) and the output gap (vertical axis). The large point labeled “Actual” represents the actual RMSs (computed in step 2); the large point labeled “Zero Shocks” is the RMS from the zero-shock counterfactual simulation (step 3); and the small points are the RMSs from the random simulations (step 4). The diagram also presents three feasible frontiers (marked by F) of the gaps’ RMSs—limited by once $F(1X\sigma^r)$, twice $F(2X\sigma^r)$, and three times $F(3X\sigma^r)$ the sample average of within-year interest rate standard deviation. In some sense, this introduces a third dimension to the diagram.

The diagram provides the following observations:

- The diagram indicates whether actual policy shocks led to outcomes that are closer to or farther away from the frontier, compared with the benchmark of zero policy shocks, and whether they reduced the RMSs of the inflation gap or the output gap.
- If the actual result is closer to the frontiers, compared with the counterfactual scenario of the zero-policy alternative, the diagram can indicate whether it reflects a liberal or a conservative attitude of the central bank.

⁹In both procedures (the counterfactual simulations and the frontier search), we discard draws leading to non-feasible negative interest rate paths. Nevertheless, we can report that for most evaluated years, this limitation is not binding (less than 0.5 percent of the randomly drawn draws are discarded). Only for the years 2009–10, during which the actual interest rate was approaching the zero lower bound, is this limitation binding, leading to a 5–25 percent discard rate. Since our (log-linear) model does not address the non-linearity associated with the zero lower bound, we take caution when interpreting the results for those years.

- The distinguished frontiers illustrate the quantitative implication of a lesser degree of interest rate smoothness.
- The diagram indicates whether the actual (smoothed) policy shocks delivered RMS results inside or outside the cluster cloud (generated by alternative policy shocks, randomly drawn around the counterfactual case of zero policy shock). In some sense, this indicates whether the actual policy actions were “significantly” different (for better or for worse) from the alternative of strictly following the monetary policy rule.

The second type of diagram (see, for example, panel B in figure 4) depicts the interest rate paths that generate the results along the frontier (the dashed lines). To economize on space, we only present the diagram related to the frontier limited by twice the sample interest volatility $F(2X\sigma^r)$. For comparison, the diagram also shows the actual path (the solid line) and the zero-policy-shock path (the line with the circle marker). This type of diagram reveals whether the actual path chosen in the evaluated period was above, below, or within the mass of efficient paths (associated with certain constraints on interest rate volatility).

For each evaluated year, we also report the within-year standard deviation of the interest rate, compared with its sample average (see, for example, the table below panel A in figure 4).

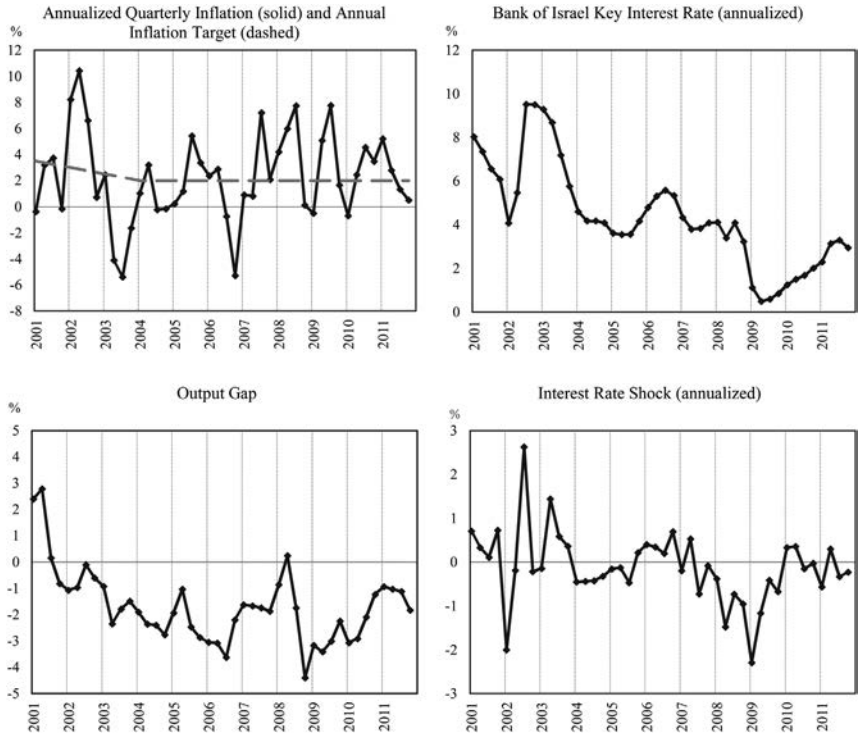
3. Implementation

This section presents an implementation of the evaluation framework to assess monetary policy in Israel during the first eleven years of the present millennium.

3.1 *Historical Perspective*

Before implementing the evaluation framework for specific years, it is worthwhile to briefly review the developments in the variables of interest during the relevant period, 2001–11. Figure 2 depicts the quarterly inflation rate, the output gap, the nominal interest rate, and the smoothed monetary policy shock consistent with equation (1).

Figure 2. The Main Variables of Interest, 2001–11



Inflation fluctuated around the midpoint of the target range, though with significant variance. Thus, the inflation rate was above target in the years 2002, 2005, 2007, 2008, 2009, 2010, and 2011 and below target in the years 2003, 2004, and 2006. The output gap was negative for most of the evaluated years, 2002–11. The recessionary downward trend of the gap during 2001–3 was due to a combination of three driving forces: a deteriorating security situation (the second “Intifada” which began in late 2000), the dot-com crash of 2000–1 (which had a large negative effect on the high-tech sector of the Israeli economy), and the impaired credibility of macroeconomic policy. The gap narrowed until mid-2008, when the Global Financial Crisis began, and started improving again in mid-2010.

The interest rate path is characterized by a downward trend, in part due to the disinflation process (see the top-left panel in figure 2)

and in part due to the reduction in long-term real yields. However, for short horizons (up to one year) variance decomposition of the nominal interest rate, as presented in Argov et al. (2012), shows the important role of the monetary policy shock: 56 percent for one quarter and 17 percent for four quarters. Inspecting the smoothed monetary policy shock (bottom-right panel in figure 2) shows three outliers: (i) In early 2002, the interest rate was unexpectedly reduced by 2 percentage points (as part of an agreement with the government on expenditure cuts), while in the second part of that year, against the background of rising inflation, there was a sharp increase in the interest rate, once again through large policy shocks. (ii) The years 2005–6 were characterized by small but persistent positive policy shocks. This occurred in the aftermath of a large depreciation in 2005 which generated concern regarding a negative interest rate gap relative to the federal funds rate. The result was low inflation and output that year. (iii) The years 2008–9 were characterized by strong and persistent negative policy shocks, reflecting preemptive measures against the background of the Global Financial Crisis.

Before we turn to the evaluation of policy for some specific years, a note on the data used in the evaluation process is in order. The data used in this work are the updated data available at the time the analysis was conducted (April 2014). To the extent that some of the data (particularly National Accounts data) is prone to revisions, the issue of what data vintage one should use when evaluating past policy decisions may arise. Note also that aside from the issue of possible data revisions, the question arises also regarding the use of future data to infer about previous periods when applying the Kalman smoother. In our view, using the most up-to-date data seems appropriate given the *ex post* nature of the analysis, where one is interested in examining policy decisions in light of our best understanding of the actual, realized circumstances *ex post*. Note that if the focus is on the *ex ante* perspective for some past period, then it would be more appropriate to use the data that was available at that time—not only when applying the Kalman smoother but also for estimating the model. Of course, this would come at the cost of losing the information embedded in the data of the succeeding periods.

We implemented our evaluation framework for each of the years 2001–11. Subsections 3.2–3.4 elaborate on a few distinct years

(presented in figures 3 to 5). In the concluding section we will provide some general observations from the full-decade evaluation.

3.2 Evaluation of 2003: Policy Shocks Increased the RMSs of Both Objective Variables

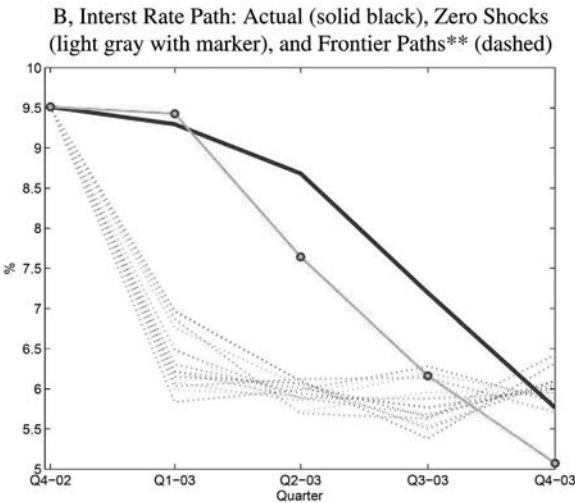
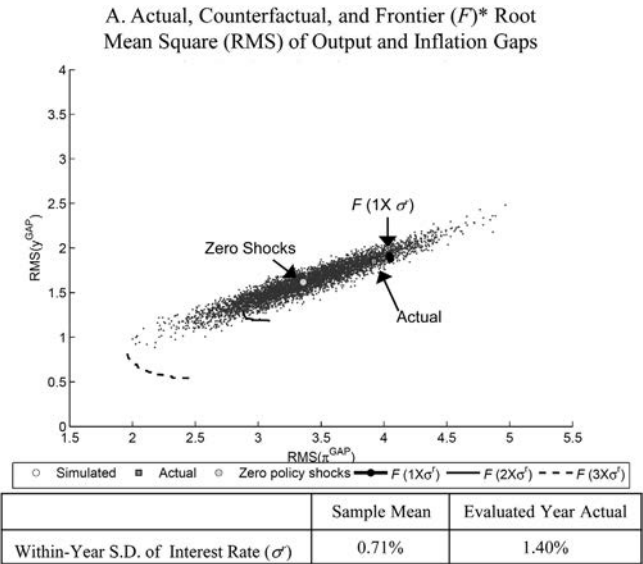
The year 2003 followed a period of stagflation that ended with a sharp rise in the interest rate to 9 percent in mid-2002. As a result, inflation declined in 2003 and fell well below the target during the second part of that year. In addition, slow growth led to a widening of the already negative output gap, which also contributed to a low inflation rate. After the realization that second-quarter inflation was too low, the interest rate was reduced to approximately 6 percent by the end of the year. According to the model, this reduction was smaller than required by the interest rate rule, implying that monetary policy shocks were in fact positive (see figure 2).

Figure 3 presents simulation results for the policy evaluation of 2003. Panel A shows 2003 to be a year in which the actual policy outcome (represented by the dark point labeled “Actual”) was dominated by that of the zero-shocks counterfactual scenario (the light gray point labeled “Zero Shocks”). Thus, the shocks actually chosen by policymakers were not in the desired direction.

We also see that the actual outcome—in terms of inflation and output volatilities—is somewhat better than those consistent with the frontier that limits interest rate volatility to the sample volatility $F(1X\sigma^r)$. However, the table below panel A reveals that the within-year interest rate volatility was actually twice as much as that of the sample. When allowing for such volatility, the RMSs could be reduced, as can be learned by the location of the relevant frontier $F(2X\sigma^r)$. Allowing for even greater variability enables greater improvement in both the inflation and output gaps’ RMS. However, these could only be generated by “abnormal” monetary policy shocks, as we see that the $F(3X\sigma^r)$ frontier is well outside the cluster cloud of randomly drawn monetary policy shocks.

It is also evident from the diagram that the frontiers are extremely short and that the cluster cloud is narrow and rising from left to right. This is a result of the low level of inflation throughout most of the year and the negative output gap. In such circumstances, lower interest rate paths would be more efficient, as they reduce the

Figure 3. Ex Post Monetary Policy Evaluation for 2003



* The figure includes three frontiers where the within-year interest rate standard deviations are limited by once, twice, and three times the sample mean.
** Frontier paths are limited by twice the sample mean of the within-year interest rate standard deviation ($2x\sigma$).

RMSs of both inflation and output gaps at the same time. Thus, in this year the policy trade-off was between jointly stabilizing the inflation and output gaps and keeping interest rate variability low. That is, the policy dilemma was how much interest rate volatility should be allowed in the process of reducing the interest rate for the benefit of higher inflation and activity.

Panel B of figure 3 indicates that the frontier-consistent interest rate paths are such that the interest rate is sharply reduced to approximately 6 percent in the first quarter of the year and remains at this level thereafter, throughout the year. This shape of paths reaches the lowest possible (average) level of interest rate given the imposed limitation on interest rate volatility (twice the sample volatility). Note that the actual interest rate was also reduced to 6 percent, but only later, by the end of the year.

The ex post conclusion is therefore that the interest rate path in 2003 was sub-optimal. It can only be seen as inefficient ex ante as well, if the low inflation from the second quarter onwards could have been foreseen. Indeed, although inflation was brought down by large nominal exchange rate appreciations, which are in general hard to predict, inflation expectations (derived from financial market data) were already below target in the second quarter of the year, suggesting that the central bank might have started to react somewhat too late, even from an ex ante perspective. As for the degree of smoothing, the somewhat traumatic experience of the sharp interest rate cut in 2002 was certainly on the minds of policymakers. Notwithstanding, we have seen that the evaluation framework only suggests that the cuts should have been made earlier, not larger. Note that the above analysis is not intended to criticize policy decisions in 2003. We are merely demonstrating that the framework generates an ex post picture of a seemingly sub-optimal policy, which can trigger important discussions on key policymaking issues, such as the implications of interest rate smoothing and the quality of inflation forecasting.

3.3 Evaluation of 2004: Efficient Policy Shocks Generated Absolute Improvement

After the relatively dramatic macroeconomic events in 2002–3, the year 2004 brought some stability. Thus, inflation was positive, for

the most part, and the negative output gap stopped widening in the second quarter. By the second quarter of the year, the interest rate had reached 4 percent, which is below the model's long-run rate. This was partly the result of negative policy shocks (see figure 2).

Figure 4 presents the policy evaluation for 2004. Panel B indicates that, similar to the case of 2003, the frontier-consistent policy paths (limited by twice the sample interest rate volatility) were generally lower than the actual one, suggesting that the reduction of the interest rate should have continued throughout the year to approximately 3 percent by the last quarter. In contrast to 2003, however, panel A shows that policy shocks in 2004 brought about improvement in the RMS of both the output and inflation gaps relative to the zero-shock alternative. Hence, this is an example of a year in which active policy decisions were in the desired direction, in terms of output and inflation stability. Yet, the frontiers demonstrate that, from an ex post point of view, greater RMS improvement could have been achieved at the cost of higher interest rate volatility.

3.4 Evaluation of 2011: Trade-Off between Inflation and Output Stabilization

The year 2011 was characterized by a relatively high rate of inflation in the beginning of the year, mainly due to a large increase in housing prices, and a negative output gap in the aftermath of the Global Financial Crisis (though narrower than the previous year). Monetary policy was characterized by interest rate hikes, which seem consistent with the interest rate rule, along with some relatively small negative monetary policy shocks (see figure 2).

Figure 5 presents the results for policy evaluation of 2011. It is evident from panel A that the actual RMSs are almost identical to the zero-shocks benchmark (as expected, considering the small monetary policy shocks). It is interesting to see that all three frontiers are virtually overlapping and lie on the border of the cluster cloud of random draws. This indicates that neither "abnormal" shocks nor increased interest rate volatility were necessary in order to reduce the RMSs of inflation and output stabilization. Moreover, we can infer that the actual outcome (i.e., inflation and output RMSs) could have been improved without significantly increasing interest rate volatility (since the actual interest rate standard deviation in that year,

Figure 4. Ex Post Monetary Policy Evaluation for 2004

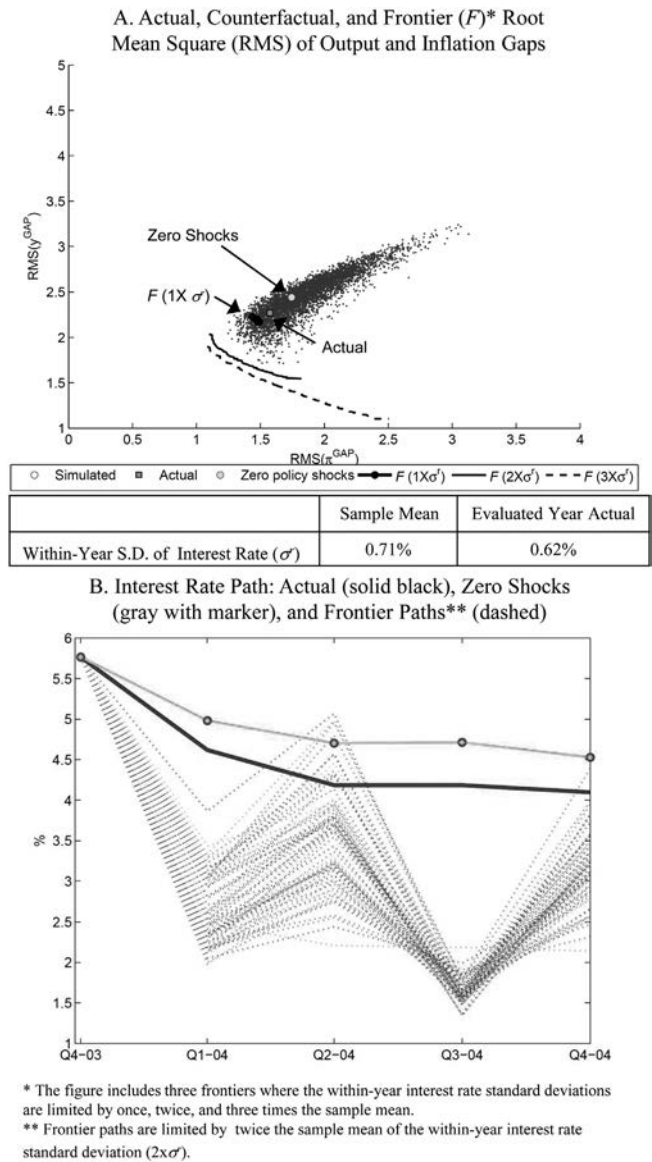
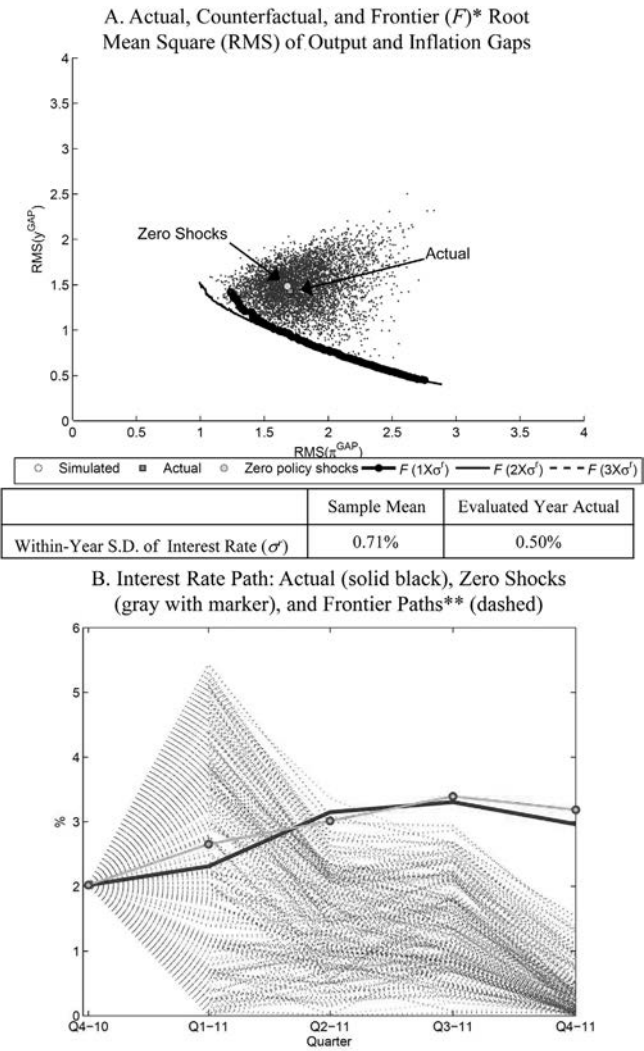


Figure 5. Ex Post Monetary Policy Evaluation for 2011



* The figure includes three frontiers where the within-year interest rate standard deviations are limited by once, twice, and three times the sample mean.
** Frontier paths are limited by twice the sample mean of the within-year interest rate standard deviation ($2x\sigma$).

at 0.5 percent, was slightly lower than the sample average, 0.7 percent, for which the frontier is drawn). While there was no trade-off between the desire for a low interest rate volatility and the other policy objectives, we can see that there was a clear trade-off between inflation and output-gap stabilization (as reflected by the slope of the frontiers). From panel B we can see that, according to one tail of the distribution of efficient interest rate paths, the interest rate should have been reduced in the beginning of the year and remained low (supporting the stabilization of the output gap) and, according to the other tail, it should have been raised in the first quarter and reduced thereafter (serving the stabilization of inflation). In practice, and in accordance with the interest rate rule, the interest rate was gradually raised throughout the year.

4. Robustness Analysis

As in any econometric evaluation, there is some uncertainty concerning the results and the conclusions. In the context of the evaluation presented in the previous section, important sources of uncertainty are the actual historical shocks and the output-gap objective of the policymaker—both unobserved to the researcher. This section examines the sensitivity of the conclusions in our implementation with respect to these two sources of uncertainty.¹⁰ In general, the results from the implementation presented above seem to be fairly robust to both sources of uncertainty.

4.1 Sensitivity to the Historical Shocks Extracted

When implementing the suggested evaluation framework, we applied the Kalman filter in order to smooth out the unobservable historical shocks. However, since the model features more shocks than observed variables, the shocks are not identified. That is, there are infinitely many combinations of shocks that are consistent with the observed variables. The Kalman filter smooths out the mean of the joint distribution of the shocks (conditional on the model, its parameters, and the full sample of the observed variables). In this section we

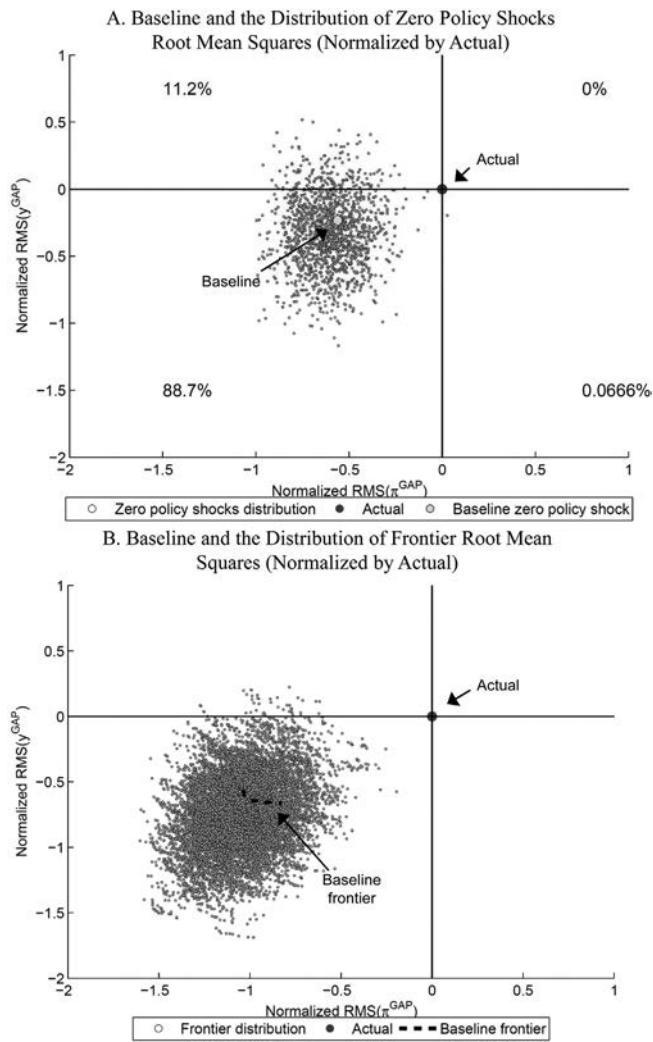
¹⁰The structure of the economy—namely, the model and its parameter values—is an additional source of uncertainty.

examine the sensitivity of the results of our implementation to the uncertainty presented by this distribution. To perform the test for a specific evaluated year, we first drew 1,500 sets of shocks from the conditional distribution.¹¹ For each set of shocks, we computed the associated RMSs of inflation and the output gap (which we refer to below as “actual”), the zero-policy-shock RMSs, and the frontier (for which we used 3,000 additional simulations). Note that although the observed variables are identical across all 1,500 iterations, the associated RMSs may nevertheless vary, as the output gap is unobservable and due to the inclusion of a forecast period in the RMS calculation (see equations (4) and (5)). Note that while the absolute location of each “actual” RMS point is uncertain, there may be less uncertainty regarding its relative location—with respect to the RMS point associated with zero policy shocks or the frontier of RMS points—which is what is relevant for our conclusions.

Figures 6–8 present results from the sensitivity analysis for the evaluated years 2003, 2004, and 2011. Each figure presents the distribution of zero-policy-shock RMSs and frontier RMSs points, all relative to the “actual” RMSs of each iteration. Thus, we “normalize” each counterfactual RMS point by subtracting from it the corresponding “actual” RMS point, so that the origin of the figure represents all the “actual” RMS points. By so shifting the location of the RMS points, we are better able to visually present our robustness analysis results, while focusing attention on the RMS points’ relative location, as discussed above. To illustrate the analysis, we will discuss the robustness test for the 2004 policy evaluation depicted in figure 7. Panel A allows examining the significance of the difference between the RMSs associated with zero policy shocks and those associated with the “actual” policy shocks. For the 2004 evaluation, we find that 74.4 percent of the normalized zero-policy-shock points are in the upper-right quadrant, implying that the absolute improvement brought about by active monetary policy is fairly robust to the uncertainty associated with historical shocks. An even more robust result is that the policy shocks did not cause an absolute worsening, which is evident in the result that only 0.27 percent of the normalized

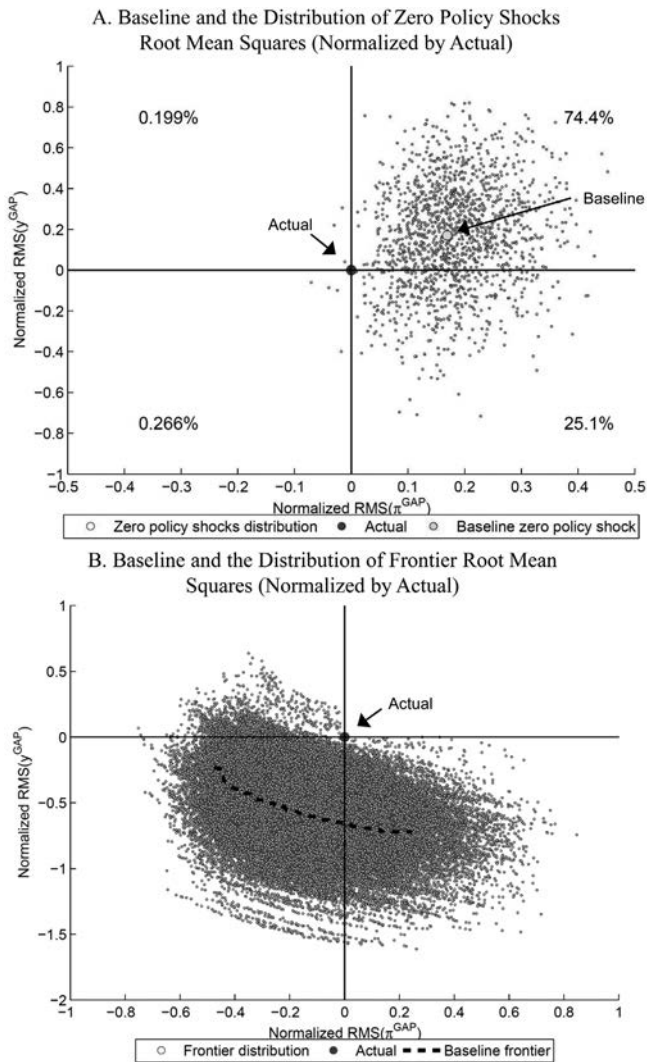
¹¹The shocks are drawn using the algorithm of Waggoner and Zha (1999), where we take historical values of observed variables to be (what they refer to as) hard conditioning restrictions. Alternatively, one could use the equivalent simulation smoother of Durbin and Koopman (2002).

Figure 6. Sensitivity Analysis for 2003 Ex Post Evaluation



RMS points are located in the bottom-left quadrant. Note that we are much more confident regarding the improvement in the inflation RMS (as 99.5 percent of the points are in the right-hand half) than for the output gap (as 25.4 percent are in the lower half). We further discuss this point below. In panel B we can see the normalized

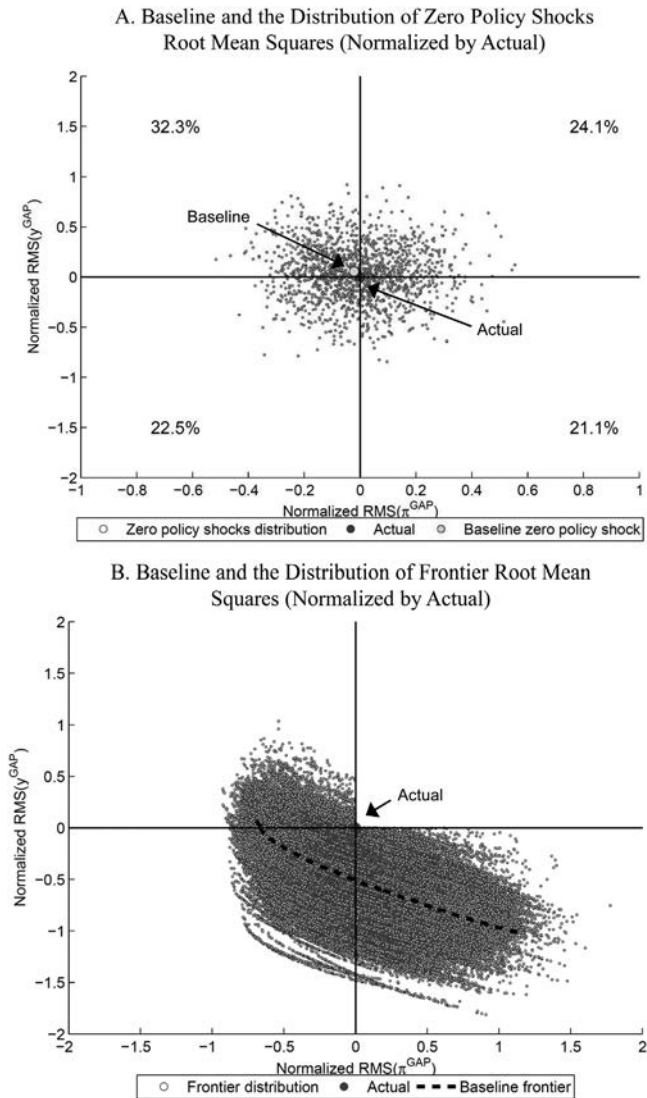
Figure 7. Sensitivity Analysis for 2004 Ex Post Evaluation



location of all $F(2X\sigma^r)$ frontier points.¹² The figure helps us to assess the uncertainty regarding the distance of the “actual” RMS from the

¹²For the sake of brevity, we only report the analysis with the frontiers that limit the within-year interest rate volatility to twice the sample average.

Figure 8. Sensitivity Analysis for 2011 Ex Post Evaluation



frontier. However, unlike the case of zero policy shocks discussed above, we cannot quantify the statistical significance of the results, since any quantitative definition for “being close to the frontier” would be arbitrary. Still, the graph shows that the mass of frontier

points is located some distance from the origin (that is, from the “actual” RMS). This result supports the conclusion that the chosen policy shocks in 2004 did not push the RMSs all the way down to the frontier. This conclusion is even more strongly evident in panel B of figure 6 for the year 2003. In contrast, the results for 2011 (figure 8), seem to be consistent with the hypothesis that actual RMSs were in the environment of the frontier.

After demonstrating the sensitivity analysis concerning the shocks smoothing, let us briefly review the factors that affect the magnitude of this uncertainty. The first type of factor pertains to the properties of the estimated model: under-identification of the shocks, as well as parameter uncertainty, induces an uncertainty with respect to the historical realizations of the shocks, which affect the robustness of the evaluation results.¹³ The second type of factor concerns the policy objectives in the evaluation framework. The “actual” state of unobserved variables (such as the output gap) depends on the smoothed historical shocks. Therefore, the inclusion of an unobservable variable in the policy criterion increases the uncertainty surrounding the evaluation result. This is evident in panel A of figures 6 and 7: the uncertainty regarding the (qualitative) position of the zero-policy-shock point relative to the actual RMS point is concentrated on the relative location of the output-gap RMS, whereas there is almost no uncertainty with respect to the direction of change of the inflation-gap RMS. Finally, an additional factor is the inclusion of the forecast period in the RMS calculation. This introduces uncertainty concerning the location of the “actual” RMS, even if both objective variables are observable. This may increase the uncertainty regarding the relative location of RMS points. To conclude, when applying the evaluation framework—to a specific country, model, or period—it is important to keep in mind these sources of uncertainty and try to address them by the relevant sensitivity tests.

¹³Our analysis is based on deviations from an estimated policy rule. In order to address the concern that the policy rule might have changed during the sample period, we reestimated the policy rule over different subsample periods. We haven't found any evidence for substantial changes in the policy rule's parameters over time.

4.2 Sensitivity to the Definition of the Output Gap

There are various alternative definitions or measures for the unobservable “output gap.” In the presented implementation, we defined the output gap as the deviation of the actual output from an unobserved, technology-driven trend (which consists of both permanent and transitory technology shocks). In this subsection we examine the robustness of our results with respect to the following alternative definitions of the output gap:

- *Baseline:* Output’s deviation from a technology-driven trend, including both permanent and transitory technology shocks ($\log \frac{Y_t}{Z_t \varepsilon_t} - \log y$). This is the measure we employed for the analysis presented above.
- *Alternative Output Gap:* Output’s deviation from a smoothed technology-driven trend, including only the permanent technology shocks ($\log \frac{Y_t}{\bar{Z}_t} - \log y$).
- *Unemployment Gap:* This measure is derived from the way Gali, Smets, and Wouters (2012) introduce unemployment into a New Keynesian model. The unemployment gap is defined as the gap between the amount of labor input that households would like to supply under the prevailing wage and the actual amount of labor input. Since unemployment is countercyclical, we take the unemployment gap with a negative sign.

Figure 9 presents the estimated historical paths of the three alternative measures. The figure reveals that the three estimates are generally similar, with the exception of only a few periods in which the signs of the measures differ.¹⁴ Focusing on the relevant time period of our ex post analysis (2001–11), we can see that, according to all three alternative measures, the output gap was negative most of the time, in line with the widespread view of the business cycles in Israel, as was noted in subsection 3.1.

Table 1 summarizes the results of some sensitivity tests conducted in order to assess the robustness of the main results in the

¹⁴Note that the sign of the output-gap measure determines whether the interest rate set by monetary policy should be higher or lower in order to improve the output gap.

Figure 9. Alternative Output Measures

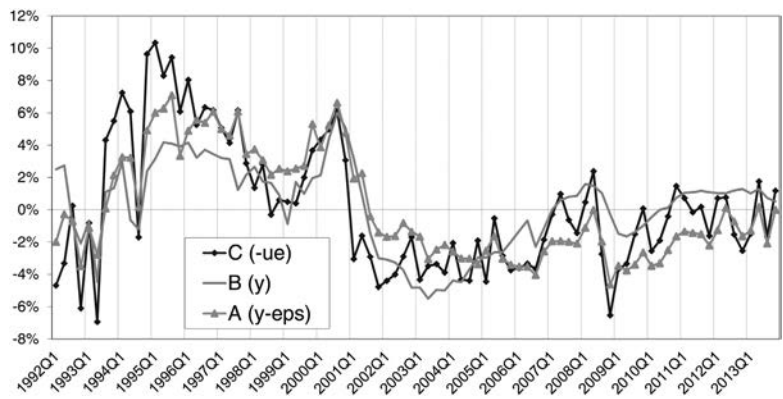


Table 1. The Main Results under Different Output Measures

	Actual-ZS			Q above Frontier			Q below Frontier		
Year	A	B	C	A	B	C	A	B	C
2001	?	+	+	1	3	3	0	0	1
2002	?	+	?	1	1	1	0	0	1
2003	+	+	+	3	3	3	0	0	0
2004	–	–	–	2	2	2	0	0	0
2005	?	?	?	2	1	1	0	0	0
2006	+	+	+	4	1	4	0	0	0
2007	?	?	?	1	1	1	0	0	0
2008	–	–	–	1	1	1	0	2	0
2009	?	?	?	1	0	1	1	1	1
2010	?	–	?	1	1	1	0	3	2
2011	?	?	?	2	1	1	0	1	1

implementation with respect to the different measures of the output gap. The block titled “Actual-ZS” describes the position of the actual outcome relative to the zero shocks (counterfactual) outcome, in the following way:

- The notation “–” denotes that the RMS of both the inflation and the output gap is smaller for the actual RMS outcome (i.e., the policy yielded an absolute improvement).
- The notation “+” denotes that the RMS of both the inflation and the output gap is larger for the actual RMS outcome (i.e., the policy yielded an absolute deterioration).
- The notation “?” denotes that one of the RMS indices has improved, while the other has deteriorated, so that none of the points dominates the other.

The table shows that except for the years 2001, 2002, and 2010, the position of the actual outcome relative to the zero-shocks outcome does not depend on the specific measure of the output gap.

The two other blocks in the table examine the robustness of the conclusions derived from the figures that compare the actual interest rate path with the interest rate paths on the efficient frontier.¹⁵ The block titled “Q above Frontier” counts, for each output measure, the number of quarters during the year in which the interest rate was higher than all of the efficient interest rate paths. Similarly, the block “Q below Frontier” counts the quarters where the interest rate was lower than all of the efficient interest rate paths. In this aspect too, the table shows that despite some differences in a few years, most of the conclusions are robust to the alternative output measures.

5. Concluding Remarks

We have presented an analytical framework for conducting an ex post evaluation of monetary policy decisions. The evaluation is based on a comparison between actual and counterfactual root mean squares (RMSs) of two objective variables (typically the inflation and output gaps) during a given period. The counterfactual RMSs are generated using counterfactual interest rate paths, calculated by drawing alternative monetary policy shocks. In addition, we search for monetary policy shocks that would have generated RMSs located on the policy frontiers, that is, where it is impossible to further improve both RMSs simultaneously. In line with the view that

¹⁵Here we refer to the frontier that is limited to twice the sample average of within-year interest rate standard deviation $F(2X\sigma^r)$.

central banks also care about interest rate smoothness, we examine different frontiers, each consistent with a different constraint on interest rate volatility.

A natural benchmark is the counterfactual case of zero policy shocks, i.e., a hypothetical case in which the central bank strictly follows the interest rate rule. Policy decisions are considered as yielding an absolute improvement (deterioration) when actual RMSs of both objective variables are lower (greater), compared with those resulting from this counterfactual benchmark scenario of zero policy shocks. We regard policy decisions as *ex post* efficient when actual RMSs are close to the frontier.

As an implementation, the proposed evaluation framework was applied to the Israeli economy for selected years. The results allow the categorizing of the evaluated years according to the following criteria: years characterized by RMSs which were close to the efficient frontier (e.g., 2011) or far away from it (e.g., 2003 and 2004); years where monetary policy shocks led to an absolute improvement—that is, to reduction in the RMS of both the inflation and output gaps—(e.g., 2004) or an absolute worsening (e.g., 2003); and years during which there was a policy trade-off between all three objective variables, including interest rate volatility (e.g., 2004), or within a subset of the objectives (e.g., 2003 and 2011).

The suggested framework has some clear strengths: it can be applied to many models, it provides flexibility in choosing the main policy objective variables and the period of evaluation, and it does not require the assessor to take a position with regard to the weights in the central bank's loss function. However, there are some weaknesses worth noting as well, which are important to keep in mind.

First, the results depend on the model being employed. To some extent, the issue of possible misspecification is common to almost any economic model. It is therefore preferable to apply the evaluation framework using more than one model (if possible). As noted in the introduction, to qualify for this purpose, models should include the objective variables of interest, as well as allow for the identification of shocks hitting the economy—in particular, shocks to the interest rate rule. To demonstrate an implementation of the evaluation framework, we applied it using the DSGE model employed by Bank of Israel's staff, MOISE. The paper describing this model, Argov et al. (2012), presents the model's properties and the way it

interprets the data. This includes the model's impulse response functions, forecast-error-variance decomposition, autocorrelations and cross-correlations (second moments) generated by the model (compared with those observed in the data), and an evaluation of in-sample forecasting quality. It concludes that the model does a fairly good job of accounting for the economic developments in the Israeli economy during the sample period. Hence, the model has been employed at the Bank of Israel as a tool in the process of forecasting, as well as for policy evaluation using the evaluation framework suggested by the present paper.

A second issue to keep in mind is that since shocks are unpredictable, the ex post approach cannot address the question of whether policy could have been more efficient in an ex ante sense. Thus, policymakers are judged (partly) based on how they reacted (in advance) to possibly unpredictable shocks. Hence, results should be interpreted accordingly. It should also be noted that the results may be sensitive to the choice of objective variables examined.

Finally, the evaluation framework views the error term of the interest rate equation as an active monetary policy shock, a view which is open to criticism. Admittedly, the error term may also represent real-time measurement errors in the variables that drive the policy rule, or errors resulting from a misspecified rule.

All these issues should be kept in mind when employing the proposed evaluation framework, and they call for the development of additional frameworks and methodologies for policy evaluation.

We believe that an evaluation on a regular basis, say once a year, may help central banks improve their understanding of past policy decisions. It may also assist central banks in communicating their decisions, and their consequences, and identifying weaknesses in the analyses and discussions preceding policy decisions.

Appendix

Search Algorithm for Frontier Investigation

In the counterfactual simulations for an evaluated period, we repeatedly drew the four-period vector of monetary policy shocks (η^R) from a zero-mean distribution $N(0, \sigma_{\eta^R}^2)$, as described in subsection 2.3 (step 4). But while searching for the feasible policy frontiers, we use

the following algorithm, intended to push draws toward the frontier, thus economizing computer time:

1. Set initial benchmark values—for the mean of the policy-shock distribution to draw from, $\eta_{benchmark}^R = 0$, and for the RMSs, $RMS(\pi_{benchmark}^{GAP}) = RMS(\hat{y}_{benchmark}^{GAP}) = \infty$. Also, select the total number of counterfactual simulations, F , and initialize the simulation number, $f = 1$.
2. Draw shock values for the policy-shock vector, η_f^R , using the distribution $N(\eta_{benchmark}^R, \sigma_{\eta^R}^2)$ for each element in the vector, and run a counterfactual simulation with this vector of policy shocks as described in subsection 2.3. Store the RMSs associated with this simulation as $RMS(\pi_f^{GAP})$ and $RMS(\hat{y}_f^{GAP})$.
3. If draw f satisfies the limit for the within-year interest rate variability, and if it falls within the desired area (defined below), update the benchmark values: $\eta_{benchmark}^R = \eta_f^R$, $RMS(\pi_{benchmark}^{GAP}) = RMS(\pi_f^{GAP})$, and $RMS(\hat{y}_{benchmark}^{GAP}) = RMS(\hat{y}_f^{GAP})$.
4. If $f < F$, then set $f \rightsquigarrow f + 1$ and go back to step 2. Otherwise, terminate.

DEFINITION 1. *Draw f is considered to fall within the “desired area” if at least one of the following conditions holds:*

- $RMS(\pi_f^{GAP}) < RMS(\pi_{benchmark}^{GAP})$ and $RMS(\hat{y}_f^{GAP}) < RMS(\hat{y}_{benchmark}^{GAP})$.
- $RMS(\pi_f^{GAP}) < RMS(\pi_{benchmark}^{GAP})$ and $(RMS(\pi_{benchmark}^{GAP}) - RMS(\pi_f^{GAP})) > \alpha * (RMS(\hat{y}_f^{GAP}) - RMS(\hat{y}_{benchmark}^{GAP}))$.
- $RMS(\hat{y}_f^{GAP}) < RMS(\hat{y}_{benchmark}^{GAP})$ and $(RMS(\hat{y}_{benchmark}^{GAP}) - RMS(\hat{y}_f^{GAP})) > (\frac{1}{\alpha}) * (RMS(\pi_f^{GAP}) - RMS(\pi_{benchmark}^{GAP}))$.

Using this algorithm, we essentially center the shocks distribution around draws that, compared with the previous draw, generate an improvement in both RMSs, or at least “a big enough” improvement in one of them. Whether an improvement is “big enough” or not depends on the parameter $\alpha > 0$, which is calibrated as the frontier slope, based on preliminary simulations.

References

- Adolfson, M., S. Laséen, J. Lindé, and M. Villani. 2007. “Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-Through.” *Journal of International Economics* 72 (2): 481–511.
- . 2008. “Evaluating an Estimated New Keynesian Small Open Economy Model.” *Journal of Economic Dynamics and Control* 32 (8): 2690–2721.
- Argov, E., A. Barnea, A. Binyamini, E. Borenstein, D. Elkayam, and I. Rozenshtrom. 2012. “MOISE: A DSGE Model for the Israeli Economy.” Discussion Paper No. 2012.06, Bank of Israel.
- Beneš, J., A. Binning, M. Fukac, K. Lees, and T. Matheson. 2009. *K.I.T.T.: Kiwi Inflation Targeting Technology*. Reserve Bank of New Zealand.
- Bernanke, B., and F. Mishkin. 1992. “Central Bank Behavior and the Strategy of Monetary Policy: Observations from Six Industrialized Countries.” In *NBER Macroeconomics Annual 1992*, Vol. 7, ed. O. J. Blanchard and S. Fischer, 183–238. National Bureau of Economic Research, Inc.
- Brubakk, L., T. Husebo, J. Maih, K. Olsen, and M. Ostnor. 2006. “Finding NEMO: Documentation of the Norwegian Economy Model.” Staff Memo No. 2006/6, Norges Bank.
- 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.
- Durbin, J., and S. J. Koopman. 2002. “A Simple and Efficient Simulation Smoother for State Space Time Series Analysis.” *Biometrika* 89 (3): 603–15.
- Gali, J., F. Smets, and R. Wouters. 2012. “Unemployment in an Estimated New Keynesian Model.” *NBER Macroeconomics*

- Annual 2011*, Vol. 26, ed. D. Acemoglu and M. Woodford, 329–60.
- Murchison, S., and A. Rennison. 2006. “ToTEM: The Bank of Canada’s New Quarterly Projection Model.” Technical Report No. 97, Bank of Canada.
- Seneca, M. 2010. “A DSGE Model for Iceland.” Working Paper No. 50, Central Bank of Iceland.
- Svensson, L. 2005. “Monetary Policy with Judgment: Forecast Targeting.” Technical Report No. 819, University Library of Munich, Germany.
- . 2012. “Evaluating Monetary Policy.” In *The Taylor Rule and the Transformation of Monetary Policy*, ed. E. F. Koenig, R. Leeson, and G. A. Kahn, 245–74. Hoover Institution Press.
- Taylor, J. B. 2011a. “Macroeconomic Lessons from the Great Deviation.” In *NBER Macroeconomics Annual 2010*, Vol. 25, ed. D. Acemoglu and M. Woodford, 387–95. University of Chicago Press.
- . 2011b. “The Rules-Discretion Cycle in Monetary and Fiscal Policy.” *Finnish Economic Papers* 24 (2): 78–86.
- Waggoner, D. F., and T. Zha. 1999. “Conditional Forecasts in Dynamic Multivariate Models.” *Review of Economics and Statistics* 81 (4): 639–51.

Interest Rate Determination in China: Past, Present, and Future*

Dong He,^a Honglin Wang,^b and Xiangrong Yu^b

^aInternational Monetary Fund

^bHong Kong Monetary Authority

How should we think about the determination of interest rates in China after interest rate liberalization? Would effective deposit rates, lending rates, and bond yields move higher or lower? We argue that interest rates in a liberalized environment would need to be anchored by monetary policy. To achieve price and output (or employment) stabilization, the policy rate should be set close to China's equilibrium or natural rate. We sketch three preliminary approaches to estimation of the natural rate. Based on this analysis, we argue that interest rates on large deposits and short-term money-market rates would likely move higher following liberalization. The effect on effective lending rates is somewhat ambiguous, as the contestability of the banking sector and the competition from the bond markets are likely to increase. We leave the determination of the curvature of the yield curve to future research.

JEL Codes: E43, E52, O53, P24.

1. Introduction

Interest rate liberalization is an essential part of China's price reforms. While most goods and services are now market determined,

*The authors would like to thank Steve Barnett, Yi Gang, John C. Williams, and one anonymous referee, as well as participants at the conferences held at the China Finance 40 Forum in September 2013 and at the Hong Kong Institute for Monetary Research in January 2014 for their comments on the paper. The views expressed in this paper are those of the authors and do not necessarily reflect those of their affiliations. Author contact: dhe@imf.org (He), wanghonglin-mail@yahoo.com (Wang), and xiangrongyu@gmail.com (Yu).

there remain significant government controls on prices of the factors of production. Ceilings imposed by the People's Bank of China (PBC) on deposit rates are a notable binding constraint on price discovery in the financial markets. Even though interest rates in the money and bond markets and bank lending rates are now market determined, their levels and movements are constrained by regulated deposit rates, due to the dominance of the banking sector in overall financial intermediation. The Chinese authorities have recently announced plans to speed up the pace of interest rate liberalization, raising the expectations of a final removal of the deposit rate ceilings in the medium term. Accompanying the process will be an overhaul of the monetary policy framework, away from a focus on credit targets toward a focus on interest rates, and an expansion of the fixed-income markets, facilitating the formation of a complete risk-free yield curve.

How should we think about the determination of interest rates in China after they are fully liberalized? Would effective deposit rates, lending rates, and bond yields move higher or lower? In this paper we argue that interest rates in a liberalized environment would need to be anchored by monetary policy. To achieve price and output (or employment) stabilization, the policy rate should be set close to China's equilibrium or natural rate. We sketch three preliminary approaches to estimation of the natural rate. Based on this analysis, we argue that interest rates on large deposits and money-market rates would likely move higher following liberalization. The effect on effective lending rates is ambiguous, as the contestability of the banking sector and the competition from the bond markets are likely to increase.

At present, the ceilings on deposit rates are imposed by the PBC through administrative orders. After interest rate liberalization, the PBC will use the composition and size of its own balance sheet to influence market rates. As the monopoly supplier of high-powered money, the PBC will be in a good position to steer short-term interest rates. As an interim measure, the PBC may still need to set targets for medium-term rates and manage the shape of the yield curve through direct interventions. This is because the bond markets will take time to expand and mature.

2. Current Interest Rate Structure

In the past three decades, China has made substantial progress toward a market-oriented system of interest rates. According to Yi's (2009) account, more than 123 interest rates had been liberalized and only 24 interest rates were still regulated as of the end of 2008. At present, interest rates can be grouped into the following four types.

2.1 Interest Rates for Central Bank Operations

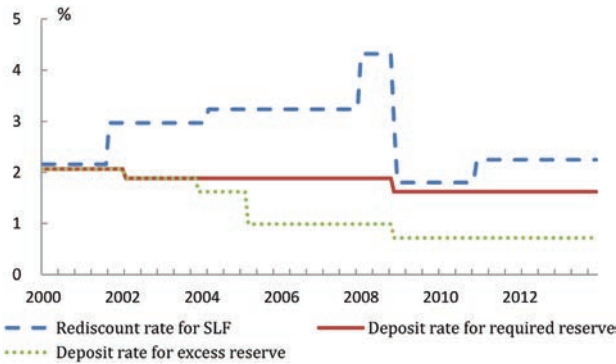
In the era of the planned economy, financial resources were allocated by government plans, and all interest rates were set by the central bank. The PBC, as the only bank in the economy, used a system of "centralized deposits and credits" to manage funds. Financial reforms in the early 1980s helped to establish a commercial banking system dominated by a number of large state-owned banks. The PBC began to function exclusively as a central bank in 1984, and a series of interest rates were created to conduct monetary and financial stability policy along with other quantitative instruments such as loan quotas and reserve requirements.

Interest rates for PBC operations can be divided into three types: first, lending rates of refinance facilities, including central bank relending rates, rediscount rates, and lending rates of the recently created Standing Lending Facility (SLF); second, deposit rates for required reserves and excess reserves; third, open-market operation (OMO) rates, including central bank repo rates, central bank bill rates, and interest rates of the newly created short-term liquidity operations (SLO). Figure 1 shows changes in rediscount rates and reserve interest rates in the last decade. In recent years, the central bank repo rates and central bank bill rates have become increasingly important tools to signal the PBC's policy intentions and to guide market interest rates toward desired levels (Zhang 2012).

2.2 Interest Rates of Banking Products

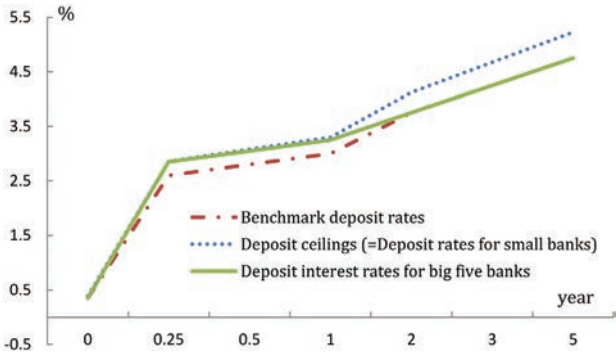
In the banking sector, most deposit rates are regulated through administratively imposed ceilings, and there were lending rate floors before they were removed in July 2013. As shown in figure 2, the

Figure 1. Interest Rates for Central Bank Operations



Source: CEIC.

Figure 2. Deposit Rates of Different Tenors



Source: CEIC.

PBC implicitly shapes the term structure by imposing ceilings on deposit rates of different maturities. These ceilings are the most powerful instrument to affect market interest rates and bank lending rates, and they help to anchor the whole interest rates system (He and Wang 2012, 2013). This is because deposits are the main funding source for commercial banks, which dominate the Chinese financial system. On the other hand, although no longer regulated, lending rates may not fully reflect market forces: First, aggregate

loan supply is in practice subject to an implicit credit target.¹ Second, lending rates are affected by regulated deposit rates owing to market frictions such as “tie-up” contracts that require borrowers to deposit their cash in the lending bank in order to obtain the credit.

Interest rates used internally by banks for capital budgeting and transfer pricing are still based on benchmark deposit rates instead of market rates such as repo rates (Jiang 2012). This further illustrates the anchor role of regulated deposit rates in the banking sector.

2.3 Interest Rates in the Money and Bond Markets

The interbank lending market in China was established in the mid-1980s. A unified trading platform was established in Shanghai in 1996 to enhance the surveillance on interbank activities, based on which the China Interbank Offered Rates (CHIBORs) formed. With more stringent regulations on uncollateralized interbank lending, the introduction of the repo market in 1997 soon replaced the interbank loan market to become the major source of short-term liquidity for banks. The rapid growth of the repo market increased the need for reference rates to price money-market instruments, resulting in the compilation of repo fixing rates for overnight, seven-day, and fourteen-day tenors. The high turnover made the corresponding fixing rates a widely used benchmark in the interbank market. However, the relatively short maturity in repo transactions has restrained interbank funding with longer tenors. The authorities also realized that a benchmark money-market yield curve would enhance the efficiency of price discovery. As a result, the Shanghai Interbank Offered Rates (SHIBORs) were introduced in January 2007, which not only increased turnover in the interbank loan market but also facilitated the pricing of a range of interest rate derivatives.

Due to the more stringent regulations on interbank lending at longer maturities, a large portion of interbank transactions are concentrated at short tenors of one month or less. Over 90 percent of transactions in the interbank market were overnight or at seven-day tenors in 2013, and more than half of the trading days registered

¹In practice, the PBC still uses moral suasion, so-called appropriate loan amount (implicit loan quota), to influence lending activities of commercial banks. For example, when this quota becomes binding for a bank at the end of the quarter, borrowers might have to pay higher interest rates to compete for loans.

no transactions for lending of six months to one year in the repo and interbank loan markets during the period 2010 to 2013. This uneven distribution has restrained use of the money-market rates as the benchmark for pricing loans with longer maturities such as corporate and syndicated loans.

The Chinese bond markets also took off in the middle of the 1980s, and the interbank bond market has grown very rapidly since the mid-2000s. However, the bond markets remain segmented between the interbank market and the exchange market, and liquidity is relatively low compared with those in advanced economies. Anecdotal evidence suggests that entry to the primary market is still somewhat restricted. Prices in the secondary market are largely market determined. However, some regulations on the pricing of bonds remain, though they may not be fully enforced. For example, the issuing rates of corporate bonds cannot be 40 percent higher than benchmark deposit rates with the same maturity. In addition, the maximum issuing size is only up to 40 percent of the issuers' net assets.

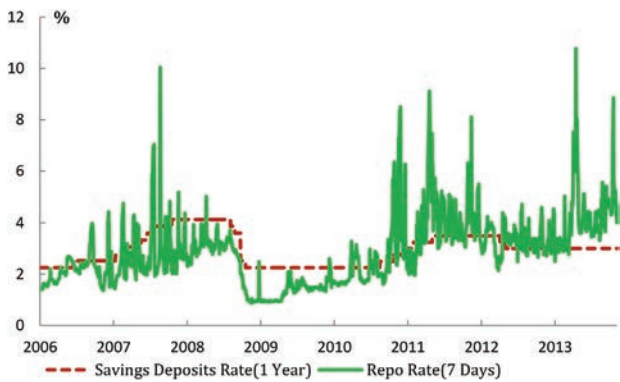
It is worth noting that observed interest rates in the money and bond markets not only reflect market forces but are also influenced by regulated bank rates. As shown in figures 3 and 4, both the seven-day repo rate and the one-year treasury yield seem to follow the pattern of the regulated one-year deposit rate.

2.4 Interest Rates in Other Financial Markets

One important recent development in the Chinese financial system is the increase in off-balance-sheet and shadow banking activities, which are incentivized by a desire to circumvent interest rate regulations and demand for alternative investment opportunities. Interest rates in the shadow banking sector are largely market determined but constrained by deposit rate regulations. Although more volatile, interest rates on wealth-management products are closely linked to interest rates on formal deposits (figure 5).

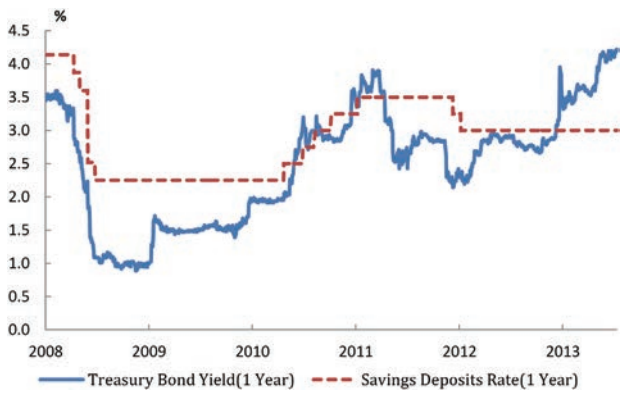
In contrast to shadow banking rates, interest rates in the informal credit markets seem much more opaque and higher (figure 6). Higher interest rates can be partially attributed to higher risk premia in the informal markets, in which major borrowers are small and medium enterprises. Despite their opacity, informal lending rates do

Figure 3. Seven-Day Repo Rate and One-Year Deposit Rate



Source: CEIC.

Figure 4. One-Year Treasury Yield and One-Year Deposit Rate



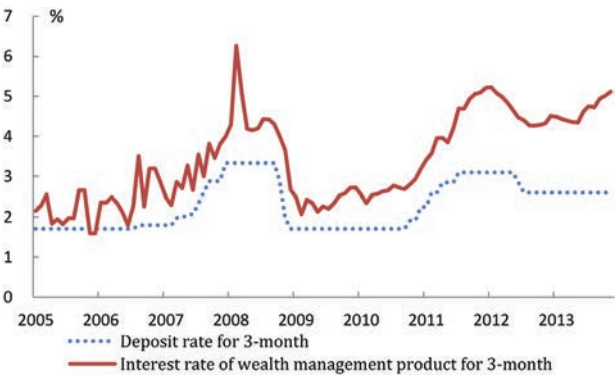
Source: CEIC and WIND.

react to changes of monetary policy and deposit rates (Qin, Xu, and Zhang 2013).

3. Rationale for Interest Rate Controls

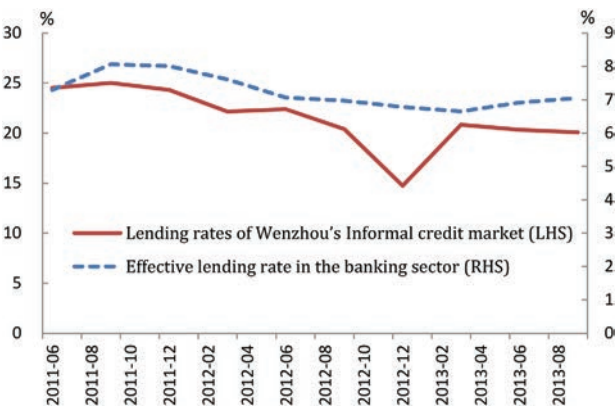
Why does China still retain control of deposit rates in the banking sector? There are a number of explanations. According to Lardy

Figure 5. Interest Rates on Wealth-Management Product and Three-Month Deposit



Source: CEIC and WIND.

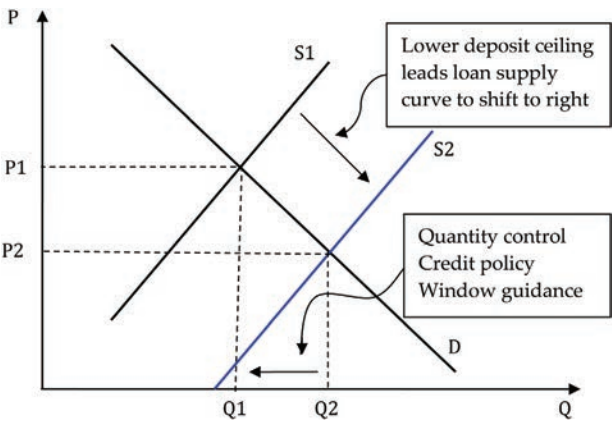
Figure 6. Lending Rates in the Informal Credit Market and Bank Lending Rate



Source: CEIC and WIND.

(2008), bank rates are kept at artificially low levels to extract transfers or subsidies from households to the corporate sector, particularly the state-owned enterprises (SOEs), and this is part and parcel of a state-led development strategy. In contrast, PBC officials tend to emphasize that interest rate liberalization requires major players

Figure 7. Credit Market with Interest Rate Distortions



Source: He and Wang (2012).

in the financial system to have hard budget constraints and be in a position to exercise self-discipline in their pricing decisions, and that the demand for credit by borrowers should also be sensitive to interest rates (Zhou 2010). However, banks and their borrowers, particularly SOEs, have not met these standards yet. In addition, when the bond markets are underdeveloped, monetary policy seems more effective through direct interest rate controls rather than through market transmission.

Regardless of policy intentions, in a situation in which the deposit rate ceilings are persistently below their equilibrium level, the PBC must rely on quantity-based instruments to achieve its policy objectives. As shown in figure 7, a lower deposit rate ceiling (compared with its equilibrium level) shifts the loan supply curve of commercial banks to the right ($S1 \rightarrow S2$, where $S1$ is the supply curve without the deposit rate ceiling). With the shifted supply curve ($S2$), banks are willing to lend at lower interest rates because their funding costs are lower than they should be ($P2 < P1$). Meanwhile, the demand for loans is also higher than its equilibrium level ($Q2 > Q1$). However, in the new equilibrium under the deposit rate ceiling ($P2, Q2$), there is more credit ($Q2 > Q1$) in the economy compared with the original equilibrium ($P1, Q1$), which may be inconsistent with the PBC's inflation target. The PBC will then have to constrain credit

supply using quantity-based instruments such as aggregate credit quotas and reserve requirement ratios. Alternatively, a lending rate floor may be used to curb loan demand through higher borrowing costs, as was the case until recently. A lending rate floor together with a deposit rate ceiling guaranteed profit margins for banks.

This simple analysis demonstrates that distortions caused by price regulations have to be corrected by quantity-based instruments. The importance of quantity-based instruments diminishes when the deposit rate ceilings are close to the equilibrium level of deposit rates. Thus, a necessary step in moving toward a primarily price-based monetary policy framework in the broader context of interest rate liberalization is for the PBC to move and maintain its benchmark or policy interest rates close to their equilibrium levels.

4. Interest Rate Determination after Full Liberalization

In a typical modern market economy, short-term rates are largely anchored by the central bank's policy rate, while longer-term rates are mainly determined by market forces, adjusting for term and risk premia. We believe that the PBC's framework would eventually converge to one like this. Although this process may take time, a useful first step is to understand the factors that would shape the PBC's analytical thinking in determining its own policy rate.

Operationally, a fundamental change is that the PBC will use the composition and size of its own balance sheet to control and influence market rates instead of using administrative others to regulate rates. As an interim measure, the PBC may consider setting targets for both short- and medium-term interest rates and managing the shape of the yield curve through direct interventions in the bond markets. This is because the bond markets will take time to expand and mature to develop an effective yield curve determined primarily by market forces. With this big picture in the mind, we discuss the formation of policy rates, bank rates, and market rates in the post-liberalization era in the rest of this section.

4.1 The Determination of Policy Interest Rates

The PBC has not yet articulated in public the contours of its monetary policy framework after interest rate liberalization. In our view,

it would likely involve setting targets for short-term interest rates through OMOs in the money markets. The Taylor rule is then a simple analytical tool for understanding the conduct of monetary policy, which specifies adjustment of the policy rate around the natural rate of interest.

The notion of the “natural,” “equilibrium,” or “neutral” rate of interest, which are sometimes used interchangeably, has been developed in the literature to provide a benchmark summarizing the economic circumstances against which we evaluate interest rates. It can be defined as the equilibrium real interest rate consistent with stable low inflation and potential output, in the absence of transitory shocks to demand and supply.² We use different methods to gauge the natural interest rate in China. The purpose is not to establish a precise number but to provide a robust starting point for future policy experiments.

4.1.1 Calibration-Based Estimates

The seminal work by Laubach and Williams (2003) estimated the natural real interest rate and potential output growth jointly by applying Kalman filtering techniques to a system of reduced-form equations that describe relations between the natural rate and observables such as output and inflation. Garnier and Wilhelmsen (2005), Manrique and Marqués (2004), and Mesonnier and Renne (2007), among others, followed a similar methodology to measure the natural rate of interest for different economies.

This method cannot be applied directly to the Chinese data. Without a structural model characterizing the influence of interest rate controls on observables, the specification applicable to advanced economies with free interest rates cannot correct for the potential institutional bias in China. Nevertheless, the critical equation in Laubach and Williams (2003) is

$$r^* = \frac{1}{\sigma}g + \theta, \quad (1)$$

²Estimation of the natural rate should correct for the effect of interest rate controls and financial constraints in China. We are interested in the natural rate after liberalization, while the actual data on interest rates and other macroeconomic variables that we observe contain systematic distortions.

**Table 1. Natural Interest Rates under
Different Calibrations**

Parameter	(1)	(2)	(3)	(4)
Intertemporal Elasticity of Substitution (σ)	2.20	2.00	2.00	1.53
Time Preference (θ)	0.3%	0.3%	0.45%	0.45%
Potential Economic Growth (g)	7.5%	7.5%	7.5%	7.5%
Natural Interest Rate	3.71%	4.05%	4.2%	5.35%
Source: Authors' calculation.				

where r^* is the natural real interest rate, σ is the intertemporal elasticity of substitution (IES), g is the growth rate of potential output, and θ is the rate of time preference. This simple relation holds in the steady state for a wide range of models. It can be derived from neo-classical models like Ramsey (1928).³ In New Keynesian models, the natural rate can be written in a similar form (Gali 2003). We may use equation (1) to infer the natural rate of interest by calibrating the parameters.

We choose parameter values that we consider reasonable for China. Following Song, Storesletten, and Zilibotti (2011), we calibrate θ to 0.3 percent to match China’s average aggregate saving rate. Given the high saving rate, research generally assumes that households in China are more “patient” than in other economies. For example, the calibration of Funke and Paetz (2013) corresponds to an average θ of 0.45 percent. For the IES, we use the value set by Song, Storesletten, and Zilibotti (2011) as a baseline, i.e., $\sigma = 2$. Garnier and Wilhelmsen (2005) estimated σ to be around 1.53 for Germany. As a similar high saving economy, we use this value as another reference. To reflect potentially higher intertemporal substitution, we also consider 2.2 for σ . For potential growth, we set g at 7.5 percent in light of the historical experience and the recent economic slowdown.

We obtain estimates for the natural interest rate in China in table 1. Columns 2 and 3 seem to be reasonable baseline estimates,

³We derive equation (1) in the context of the Ramsey model in the appendix.

while columns 1 and 4 can be considered the lower and upper bounds under our parameterization. Based on these results, the natural rate in China is about 4 percent or slightly higher.

Of course, these estimates depend on our choice of parameter values. For example, if economic growth slows to 7 percent, the natural rate will adjust downward by 0.25 percent in the baseline case.⁴ The reasonable range for the IES is also wide, and we hope that future research can provide evidence for the IES in China.

4.1.2 *Marginal Returns to Capital*

Our second approach is to measure the natural interest rate using equilibrium conditions from the production side of the economy. We consider a standard macroeconomic model featuring a constant-returns-to-scale production function. Under the assumption of competitive markets, total output is distributed among factors according to their marginal productivity, without surplus or deficit. With these minimal conditions, the marginal product of capital MPK is given by

$$MPK = \frac{\alpha}{K/Y}, \quad (2)$$

where Y is the total output, K is capital stock, and α is the share of capital income. The equilibrium interest rate r^* then equals the marginal product of capital net of tax and depreciation:

$$r^* = (1 - \tau)(MPK - \delta), \quad (3)$$

where τ is the ratio of tax on capital income and δ is the depreciation rate. All variables are expressed in real terms. The tax deduction of depreciation expenses is allowed.

The output Y is proxied by China's GDP from the National Bureau of Statistics (NBS) deflated by the GDP deflator provided by the International Monetary Fund (IMF). We use the perpetual inventory method to construct estimates of the capital stock. To be

⁴Some observers are concerned about the possibility of more dramatic slowdown, but the natural interest rate is a concept more relevant in normal situations.

consistent with the measure of the World Bank (2011), on which our method will partly rely, we proxy investment by gross capital formation deflated by the fixed-asset investment price index.⁵ The initial capital stock is given by the ratio of the first-year investment to the sum of the average growth rate of investment in the first five years and the depreciation rate, which is set at 5 percent. The rest of the capital stock series derives from the capital accumulation process, i.e., adding investment to existing stock, net of depreciation, yields current capital stock. From a macroeconomic viewpoint, capital income includes taxes on both output and enterprise income. We adjust for value-added tax, enterprise income tax, and operation tax, using the NBS data. In general, this construction of the capital stock series is standard, and the results are comparable to those of, for example, Bai, Hsieh, and Qian (2006) and Chow (1993).

In the literature, it is common to back out the capital income share α as one minus the labor share α_L . However, $1 - \alpha_L$ includes payments accruing to both produced and natural capital, i.e., land and natural resources. In contrast, the measure of capital stock is calculated from investment flows that cover only machinery, equipment, and plant and therefore only represents produced capital. Using $1 - \alpha_L$ as the capital share biases the estimate of MPK upwards. Given that agricultural and natural resources have a large share in inputs in developing China, the potential biases can be serious.

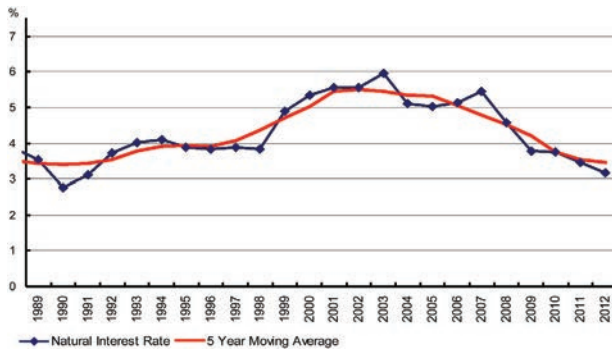
To construct α , we first calculate the labor share α_L by taking the average of the provincial labor shares weighted by GDP using data from the NBS. We then adjust for the payments to natural capital N to obtain the capital income share α corresponding to our measure of produced capital K :

$$\alpha = \frac{K}{K + N}(1 - \alpha_L). \quad (4)$$

This method was proposed by Caselli and Feyrer (2007), who documented that the returns to capital are remarkably similar across

⁵Here we use different price indices to deflate GDP and investment. In a developing country like China, investment goods are more expensive than final goods. Not accounting for these differential prices can lead to overestimation of the return to capital.

Figure 8. MPK Estimates of the Natural Interest Rate in China



Source: Authors' calculation.

countries after these necessary adjustments. The Wealth of Nations data set compiled by the World Bank (2011) contains estimates of produced and natural capital in 1995, 2000, and 2005. For China, the three observations on the share of produced capital in total capital are 0.45, 0.50, and 0.60, respectively. We take an average and assume that it prevails across time.

Figure 8 shows that the estimates of the natural real interest rate in China average around 4.3 percent, within a band between 3 percent and 6 percent. The results are relatively stable over time. Since the capital stock partially depends on current investment, which is procyclical, we also provide a moving-average measure to remove high-frequency components. It is worth pointing out that these estimates represent the net returns to capital in an aggregate sense, absorbing the effects of uncertainty and liquidity needs. In this basic non-monetary setting, there is in equilibrium only one interest rate which corresponds to the natural interest rate. Our estimation requires only minimal assumptions of the economic structure, and hence the estimates should be robust to model specification.

4.1.3 Cross-Country Experience

Following He and Wang (2012), we correct for the effects of interest rate control and financial repression on market interest rates to

obtain estimates of the natural rate. They used a panel data set of forty-nine economies between 1973 and 2005 to estimate the linear regression model that links real interest rates r_i to real GDP growth rate g_i , the aggregate saving rate θ_i , and the degree of financial repression τ_i :

$$r_i = a_0 + a_1g_i + a_2\theta_i + a_3\tau_i + \pi_i + u_i, \quad (5)$$

where π_i is the country fixed effects. The saving rate is included to capture households' time preference. The real interest rate is the lending interest rate adjusted for inflation as measured by the GDP deflator, which is obtained directly from the World Development Indicators database of the World Bank. The financial repression index is between 0 and 1, compiled by the IMF (2008). It is aggregated from liberalization scores for credit controls and reserve requirements, interest rate controls, entry barriers, state ownership, policies on securities markets, banking regulations, and restrictions on the capital account.⁶ He and Wang (2012) showed that financial repression has a significant negative effect on observed real interest rates and obtained an estimate of about 4.7 percent for China's natural rate in 2005.

Since data on financial repression are not available after 2005, we cannot update the estimation of (5). Instead, we use linear interpolation to estimate the current level of financial repression in China. Using the empirical results of He and Wang (2012), we remove the distortions caused by financial repression on China's most recent data on real interest rates. A simple moving average is employed to smooth out cyclical shocks. In this way, we find that the natural rate has declined somewhat but remained above the level of 4 percent. The most recent estimate is 4.31 percent for 2012.

Taken together, our results from the above three methods are largely consistent. A reasonable interval for the estimate of the natural interest rate is between 4 percent and 4.5 percent in China. This is higher than Taylor's (1993) 2 percent and the estimate of Laubach and Williams (2003) of around 3 percent for the U.S. economy. Our methods do not explicitly model the term structure of interest rates.

⁶See IMF (2008) for more details. The data set is available at <http://www.imf.org/external/pubs/cat/longres.aspx?sk=22485.0>.

Corresponding to the policy rate, a small downward adjustment over our estimates may be appropriate.

4.2 The Determination of Interest Rates in the Banking Sector

In a fully liberalized interest rate system, it is likely that interest rates on large deposits and negotiable certificates of deposit (CDs) will track wholesale funding costs such as money-market rates closely, whereas interest rates on small deposits would probably be somewhat lower. When market arbitrage works effectively, money-market rates and interest rates on large bank deposits and CDs are well anchored by the central bank policy rates. According to our estimates of the natural interest rate, there will be pressures for bank deposit rates to move higher when they are no longer constrained by the binding ceilings, *ceteris paribus*.⁷ Indeed, most banks raised their deposit rates close to the ceilings after the PBC announced that the deposit ceilings would be raised from 100 percent to 110 percent of the benchmark deposit rates in 2012.

The movement of lending rates after liberalization would depend on, among other things, the following three factors. The first is the contestability of the banking industry. Since the effective funding costs of the banking sector (deposits) are set to increase after liberalization, banks might be able to pass higher costs on to their clients if they have sufficient market power, resulting in higher effective lending rates that borrowers have to accept. Second, lending rates will depend on whether the PBC continues to impose an aggregate loan quota and whether the quota is binding after liberalization. If the quota remains binding, the effective loan rate would be determined by the quota and aggregate loan demand. The third factor is the competition from direct finance and foreign inflows after liberalization. If the bond markets are more developed and more firms choose bond markets to finance their projects or foreign lenders are allowed to lend to domestic firms, banks would have to lower their lending rates to compete with the new rivals.

⁷ As shown in figure 2, the benchmark rate for the one-year deposit is 3 percent in 2013, up to 110 percent of which banks can offer to depositors. The inflation rate was 2.6 percent in 2013. In this example, the real one-year deposit rate was slightly lower than 1 percent.

4.3 The Determination of Interest Rates in the Money and Bond Markets

Interest rates in the money and bond markets after liberalization would be determined by both market forces and monetary policy. As argued earlier, the conduct of monetary policy in the future would mainly involve a policy rate, which is typically a short-term money-market rate. As the monopoly supplier of high-powered money, the PBC in principle can steer the market rate to whatever level desired through its operations in the markets. Given the relatively short yield curve, interest rates in the money markets should be well anchored by monetary policy after liberalization.

In the bond markets, monetary policy would continue to have its influence, in particular, at the short end of the yield curve. According to our estimates of the natural interest rate, short-term bond rates would likely move higher after liberalization. To some extent, the recent rise in China's interbank rates is an indication of such upward adjustment pressures. For longer-term interest rates, market forces would dominate in pricing term and risk premia. The mechanism is complicated, but it is clear that policy signals can still be passed on to the longer parts of the yield curve in deep well-functioning bond markets.

As a final point, market rates and bank rates will affect each other through fund flows between the banking sector and the capital markets. This is consistent with the recent observation that market rates increased as the PBC lifted the deposit rate ceilings. Shocks from the money and bond markets can also affect bank rates (e.g., lending rates), although the impact is relatively small (He and Wang 2013).

5. Next Steps

In a narrow sense, there seems to be only one step left to achieve interest rate liberalization, which is the removal of the deposit rate ceilings. However, its success and the functioning of the financial system in serving the economy afterwards depend on many conditions. It should also be noted that the liberalization process is a dynamic one: market-determined interest rates at the margin will

help to reduce distorted incentives and moral hazard, thereby creating conditions for further reforms. China's success in other areas of price reform through the dual-track approach bodes well for its success in interest rate reforms.

A carefully designed roadmap is necessary to guide the series of reforms through the process. Governor Zhou (2013) gave his view on how to achieve interest rate liberalization in two steps. In the near term, the focus is to strengthen self-discipline in the formation of market interest rates, to grant financial institutions more discretion in setting their interest rates, to establish the prime lending rate fixing as an effective benchmark for the pricing of loan products, to promote the issuance and trading of negotiable CDs, and to gradually expand the scope of market-based pricing of the liabilities of financial institutions. In the medium term, the focus should be on developing a relatively complete and efficient market interest rate system and improving the monetary policy framework and monetary policy transmission.

The authorities are likely to adopt the following two approaches to deregulation of bank deposit rates. First, the PBC may lift the deposit rate ceilings gradually and allow deposit rates to move upwards in a managed way, and remove the ceiling regulations when actual deposit rates are close enough to their equilibrium levels. Second, the PBC may allow parts of banks' liabilities other than deposits, such as CDs issued by banks, to be priced by market forces. For example, negotiable CDs can be issued to institutional investors first and then extended to individual investors. The face value of a CD can initially be set large and gradually reduced so that the CDs resemble retail deposits more.

The success of these reforms requires at least two conditions. First, major players in the financial system, particularly banks, should have hard budget constraints and be in a position to exercise self-discipline in their pricing decisions. Second, the demand for credit by borrowers, especially SOEs and financial firms backed by local governments, should be sensitive to changes in interest rates. In particular, implicit government guarantees have to be removed for a sound banking system. Otherwise, banks may expect the government to bail out potential failures when competing for deposits by raising interest rates and/or chasing after returns by lowering lending standards. Establishing an explicit exit mechanism for financial

institutions is a priority and can contain moral hazard. A deposit-protection scheme would be an important element in developing the conditions for fully liberalized deposit rates, as suggested by the experience of many other economies. The PBC is generally expected to launch such a scheme soon.

In short, as a price discovery process, interest rate liberalization demands that banks, firms, and households are responsible for their own decisions and aware of the risks associated with any given return. Before that, the PBC may still need to use measures like aggregate loan quotas even after liberalized deposit rates have moved close to the equilibrium levels.

6. Concluding Remarks

In this paper, we have discussed interest rate determination in China before and after interest rate liberalization. There are many questions that we have not addressed. For example, we have argued that short-term interest rates in the bond markets may increase after liberalization on the basis of our empirical findings, but we have not examined the potential effects on the curvature of the yield curve. The liquidity preference and risk appetite of market participants may shift, following interest rate reforms. Regulatory and market structures may also change and shape pricing behaviors. Moreover, we have focused on a closed-economy setting throughout this paper. As a continental-sized economy and with more flexible exchange rates, China is expected to be able to conduct fairly independent monetary policy and anchor the short-term interest rates well. However, since the authorities have also announced plans to accelerate the pace of capital account liberalization, the size, direction, and volatility of gross and net capital flows would inevitably affect the interest rate structure in China. These issues are left for future research.

Appendix. Equilibrium Interest Rate in the Ramsey Model

Consider a standard neoclassical model like that of Ramsey (1928). The representative agent has an instantaneous utility function of the constant-relative-risk-aversion form:

$$u(C(t)) = \frac{C(t)^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}},$$

where $C(t)$ denotes consumption and σ governs the elasticity of substitution between consumption at any two points in time. Let L denote labor supplied by the agent and A denote the effectiveness of labor. Assume that the rate of technological progress is g , i.e., $A = gA$. The law of motion of the ratio of the agent's asset to effective labor, $x = X/AL$, is then obtained as

$$\dot{x}(t) = (r(t) - g)x(t) + w(t) - c(t).$$

Taking interest rate $r(t)$ and wage rate $w(t)$ as given, the agent chooses paths of consumption $c(t) = C(t)/AL$ and asset $x(t)$ to maximize utility at time $t = 0$:

$$B \int_0^\infty \exp(-\beta t) \frac{c(t)^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} dt,$$

where $B \equiv A(0)^{1-\frac{1}{\sigma}} L(0)$ and $\beta \equiv \rho - (1 - \frac{1}{\sigma})g > 0$, with ρ being the rate of time preference. Subject to the law of motion of $x(t)$ and the no-Ponzi-game condition, maximization yields

$$\frac{\dot{c}(t)}{c(t)} = \sigma \left[r(t) - \rho - \frac{1}{\sigma}g \right].$$

Meanwhile, firms employ the stocks of labor and capital, pay them their marginal products, and sell the resulting output. And all markets clear.

On the balanced growth path, we must have $\dot{c}(t) = 0$, which implies

$$r^* = \rho + \frac{1}{\sigma}g.$$

Since output grows at constant rate g on the balanced path, this gives equation (1).

References

- Bai, C.-E., C.-T. Hsieh, and Y. Qian. 2006. "The Return to Capital in China." *Brookings Papers on Economic Activity* 37 (2): 61–102.
- Caselli, F., and J. Feyrer. 2007. "The Marginal Product of Capital." *Quarterly Journal of Economics* 122 (2): 535–68.
- Chow, G. 1993. "Capital Formation and Economic Growth in China." *Quarterly Journal of Economics* 108 (3): 809–42.
- Funke, M., and M. Paetz. 2013. "Financial System Reforms and China's Monetary Policy Framework." Presented at the Fourth HKIMR Annual International Conference on the Chinese Economy, Hong Kong, January 10–11.
- Gali, J. 2003. "New Perspectives on Monetary Policy, Inflation, and the Business Cycle." In *Advances in Economics and Econometrics*, Vol. 3, ed. M. Dewatripont, L. Hansen, and S. Turnovsky. Cambridge, UK: Cambridge University Press.
- Garnier, J., and B.-R. Wilhelmsen. 2005. "The Natural Real Interest Rate and the Output Gap in the Euro Area: A Joint Estimation." ECB Working Paper No. 546.
- He, D., and H. Wang. 2012. "Dual-Track Interest Rates and the Conduct of Monetary Policy in China." *China Economic Review* 23 (4): 928–47.
- . 2013. "Monetary Policy and Bank Lending—Evidence from Loan-Level Data." Working Paper No. 16, Hong Kong Institute for Monetary Research.
- International Monetary Fund. 2008. "A New Database of Financial Reforms." IMF Working Paper No. 08/266.
- Jiang, X. 2012. "The Funds Transfer Pricing System of Commercial Banks in China." *Journal of Finance and Economics* (in Chinese) 8: 59–61.
- Lardy, N. 2008. "Financial Repression in China." Working Paper No. PB08-8, Peterson Institute for International Economics.
- Laubach, T., and J. C. Williams. 2003. "Measuring the Natural Rate of Interest." *Review of Economics and Statistics* 85 (4): 1063–70.

- Manrique, M., and J. M. Marqués. 2004. "An Empirical Approximation of the Natural Rate of Interest and Potential Growth." Working Paper No. 0416, Banco de España.
- Mesonnier, J.-S., and J.-P. Renne. 2007. "A Time-Varying 'Natural' Rate of Interest for the Euro Area." *European Economic Review* 51 (7): 1768–84.
- Qin, D., Z. Xu, and X. Zhang. 2013. "How Much Has Private Credit Lending Reacted to Monetary Policy in China? The Case of Wenzhou." Working Paper No. 178, Department of Economics, The School of Oriental and African Studies.
- Ramsey, F. P. 1928. "A Mathematical Theory of Saving." *Economic Journal* 38 (152): 543–59.
- Song, Z., K. Storesletten, and F. Zilibotti. 2011. "Growing Like China." *American Economic Review* 101 (1): 196–233.
- Taylor, J. B. 1993. "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy* 39: 195–214.
- World Bank. 2011. *The Changing Wealth of Nations*. Washington, DC: The World Bank.
- Yi, G. 2009. "Progress of Interest Rate Liberalization in China during the 30 Years of Reform." *Journal of Financial Research* (in Chinese) 343 (1): 1–14.
- Zhang, X. 2012. *China Monetary Policy*. Beijing, China: China Financial Publishing Housing.
- Zhou, X. 2010. "Some Thoughts about Advancing Interest Rate Liberalization in China." Available at http://www.pbc.gov.cn/publish/goutongjiaoliu/524/2012/20120112160648353655534/20120112160648353655534_.html (in Chinese).
- . 2013. "Deepen All-Round Financial Reforms and Openness, Quicken the Development of a Financial Market System." Available at <http://www.pbc.gov.cn/publish/goutongjiaoliu/524/index.html> (in Chinese).

How Frequently Should We Reestimate DSGE Models?*

Marcin Kolasa and Michał Rubaszek
Narodowy Bank Polski and Warsaw School of Economics

A common practice in policymaking institutions using DSGE models for forecasting is to reestimate them only occasionally rather than every forecasting round. In this paper we ask how such a practice affects the accuracy of DSGE model-based forecasts. To this end we use a canonical medium-sized New Keynesian model and compare how its quarterly real-time forecasts for the U.S. economy vary with the interval between consecutive reestimations. We find that updating the model parameters only once a year usually does not lead to any significant deterioration in the accuracy of point forecasts. On the other hand, there are some gains from increasing the frequency of reestimation if one is interested in the quality of density forecasts.

JEL Codes: C53, E37.

1. Introduction

Dynamic stochastic general equilibrium (DSGE) models are currently the workhorse framework in macroeconomic analyses and forecasting. Their use is widespread in policymaking institutions, and central banks in particular. One of the reasons for the popularity of DSGE models is that their forecasting performance has been found to be relatively good in comparison with standard time-series models

*We are grateful to the participants of the International Symposium on Forecasting in Seoul and the seminar at the Narodowy Bank Polski, as well as three anonymous referees for useful comments. The views expressed in this paper are those of the authors and not necessarily of the institutions they represent. Corresponding author (Kolasa): Narodowy Bank Polski, ul. Świętokrzyska 11/21, 00-919 Warsaw, Poland. Tel.: +48 22 185 2465. Fax: +48 22 585 4374. E-mail: marcin.kolasa@nbp.pl.

as well as expert judgment (see, e.g., Smets and Wouters 2003; Adolfson, Lindé, and Villani 2007; Rubaszek and Skrzypczynski 2008; Edge, Kiley, and Laforde 2010; Del Negro and Schorfheide 2012; Kolasa, Rubaszek, and Skrzypczynski 2012; Wieland and Wolters 2012).

A common practice in institutions using DSGE models for forecasting is to reestimate them only occasionally rather than every time a new forecast is produced. The main reason, applying also to other econometric models, is that reestimation complicates communication between the modelers and policymakers, as the difference between consecutive forecasts is affected both by new data release and by changes in parameter estimates. Infrequent reestimation decreases the number of forecasting rounds during which this second source of the difference has to be taken into account. Also, the policymakers can get used to certain model properties—the response of the economy to a monetary shock, for instance—and hence do not want to see them change every time a new forecast is produced. Another argument supporting this practice, though its relevance is certainly diminishing due to increasing computing power of multi-core processors, is that the estimation process of large DSGE models can be very time consuming. Needless to say, if every change in model properties needs to be documented and communicated to the policymakers, frequent reestimations eat up even more time and prevent modelers from getting involved in potentially more productive projects. The final argument is that, to the best of our knowledge, there are no clear guidelines in the literature on how frequently models should be reestimated so that the accuracy of forecasts they generate is unaffected.

In contrast to real-life applications, most of the literature investigating performance of DSGE model-based forecasts is based on out-of-sample exercises where model parameters are updated quarterly. Only a few studies reestimate the model parameters at longer intervals, e.g., every four quarters (Adolfson, Lindé, and Villani 2007; Smets and Wouters 2007; Christoffel, Coenen, and Warne 2010) or even only once, at the beginning of the evaluation sample (Giannone, Monti, and Reichlin 2010). None of the above papers discuss how infrequent model reestimation affects forecast accuracy. While it is possible that using most recent data does not necessarily improve forecasts obtained with econometric models (see, e.g., Swamy and

Schinasi 1986), one can also be concerned that obsolete parameter estimates may have non-negligible costs in terms of the quality of predictions generated for some macro-aggregates influencing the policy decisions.

To provide some guidelines for practitioners, in this study we ask the following question: how often should we update DSGE model parameters to obtain efficient predictions about the main macro-variables? To answer it, we take a canonical medium-sized New Keynesian model and compare how its quarterly real-time forecasts for the U.S. economy vary with the interval between consecutive reestimations. Our main results, based on three key macroeconomic variables (output, inflation, and the interest rate), can be summarized as follows. We find that updating the model parameters only once a year does not lead to any significant deterioration in the accuracy of point forecasts. Even though there are some gains from increasing the frequency of reestimation if one is interested in the quality of density forecasts, these gains are rather small. These general conclusions are robust to using looser priors in our benchmark model or augmenting it with financial frictions. Finally, much of the decrease in forecast precision that is observed when reestimations become less frequent is because of a shorter sample rather than due to data revisions.

The rest of this paper is organized as follows. Section 2 presents the benchmark model that we use in our investigation. Section 3 describes the settings of the forecasting contest. The main results and robustness checks are discussed in section 4. The last section concludes.

2. Model

Our investigation is based on the canonical medium-sized New Keynesian framework of Smets and Wouters (2007). It features utility-maximizing households, profit-maximizing firms, a fiscal authority financing exogenous spending with lump-sum taxes, and a central bank setting short-term interest rates according to a Taylor-like rule. The model incorporates a number of real and nominal rigidities, including habits in consumption, investment adjustment costs, time-varying capacity utilization, as well as wage and price stickiness with indexation.

The exact specification we use differs from Smets and Wouters (2007) only in that we additionally allow for trend investment-specific technological progress. This modification is aimed to account for the deviation between the average growth rate of real investment and that of other GDP components.¹ A full list of log-linearized model equations can be found in the appendix.

All estimations are done using Bayesian methods and seven standard quarterly macroeconomic variables for the United States: output, consumption, investment, wages, hours worked, inflation, and the interest rate. Full definitions and sources of the real-time data used are given in the appendix. The prior assumptions are identical to those in Smets and Wouters (2007) and are also listed in the appendix. The posterior distributions are approximated using the Metropolis-Hastings algorithm with 250,000 replications, out of which we drop the first 50,000.

3. Forecasting Contest

We compare the accuracy of forecasts generated by the DSGE model, the parameters of which are reestimated according to the following five schemes:²

- (i) Update 1Q: The model is reestimated each quarter (baseline variant).
- (ii) Update 2Q: Estimation is repeated when the data for the second or fourth quarter are available.
- (iii) Update 1Y: The model is reestimated when full-year data are available.
- (iv) Update 2Y: The model is reestimated only in even years.
- (v) Fixed: The parameters are estimated once, at the beginning of the forecast evaluation sample.

¹Smets and Wouters (2007) deal with this discrepancy in long-run trends by defining real investment as nominal investment deflated with the GDP deflator. We cannot follow this path since there is no nominal investment series in our real-time database.

²In practice, some reestimations might be carried out in response to changes in model structure or data definitions rather than according to a fixed schedule. Our alternative schemes do not account for such a possibility.

Our investigation proceeds in three steps. First, we collect quarterly real-time data (RTD) describing the U.S. economy in the period between 1966:Q1 and 2011:Q4. The data are taken from the Federal Reserve Bank of Philadelphia's Real-Time Data Set for Macroeconomists, which is described in more detail by Croushore and Stark (2001). The use of RTD enables us to control for both reasons that support frequent reestimation, which we call *sample* and *vintage* effects. To be more precise, let θ_T^v be the vector of parameter estimates based on the sample 1 : T from vintage date v . The difference between the current estimates and those done q quarters ago using the data vintage available at that time can be decomposed into

$$\theta_T^v - \theta_{T-q}^{v-q} = \underbrace{(\theta_T^v - \theta_{T-q}^v)}_{\text{sample effect}} + \underbrace{(\theta_{T-q}^v - \theta_{T-q}^{v-q})}_{\text{vintage effect}}. \quad (1)$$

Hence, our main forecasting contest that uses RTD captures both sample and vintage effects. In the case in which we use the latest available data (LAD), the difference in forecast accuracy between the five schemes is only due to the sample effect.

In the second step, for each moment of forecast formulation t from the period 1989:Q4–2011:Q3, horizon h , and forecasting scheme j , we compute point forecasts— $y_{t+h}^f|t, j$ —as well as the predictive scores $p(y_{t+h}|t, j)$. More precisely, out of 200,000 posterior draws for parameter vector θ obtained with the Metropolis-Hastings algorithm, we select $N = 5,000$ equally spaced subdraws $\theta^{(n)}$ for $n = 1, 2, \dots, N$. For each $\theta^{(n)}$ we calculate the first two moments of the predictive density $p(Y_{t+h}|t, j, \theta^{(n)})$, which has Gaussian distribution.³ Subsequently, following Geweke and Amisano (2014), the point forecast is computed as the mean of the first moments:

$$y_{t+h}^f|t, j = \frac{1}{N} \sum_{n=1}^N E(Y_{t+h}|t, j, \theta_i^{(n)}), \quad (2)$$

³For each parameter draw, the Kalman filter is used to obtain smooth estimates of unobservable model states given the data available at the time the forecast is formulated (RTD variant) or the latest available data (LAD variant).

whereas the predictive score is obtained as the average of the predictive densities evaluated at the realization y_{t+h} of Y_{t+h} :

$$p(y_{t+h}|t, j) = \frac{1}{N} \sum_{n=1}^N p(y_{t+h}|t, j, \theta_i^{(n)}). \quad (3)$$

The forecasting scheme is recursive,⁴ the evaluation sample spans from 1990:Q1 to 2011:Q4, and the maximum forecast horizon is twelve quarters. Thus, the first set of forecasts is generated for the period 1990:Q1–1992:Q4 with the model estimated on the sample spanning 1966:Q1–1989:Q4. The second set of forecasts is for the period 1990:Q2–1993:Q1 with the model estimated on the sample 1966:Q1–1990:Q1 (baseline scheme) or 1966:Q1–1989:Q4 (remaining schemes). The third set of forecasts is for the period 1990:Q3–1993:Q2 with the model estimated on the sample 1966:Q1–1990:Q2 (baseline and update 2Q) or 1966:Q1–1989:Q4 (remaining schemes), etc. Since our data set ends in 2011:Q4, the number of forecasts that we can use in evaluation ranges from seventy-seven (for twelve-quarter-ahead forecasts) to eighty-eight (for one-quarter-ahead forecasts).

In the third stage, we assess the quality of forecasts for the key three U.S. macroeconomic time series: output, inflation, and the interest rate. Given that the maximum forecast horizon is relatively long, the comparisons are for output and price levels rather than growth rates. The realizations, which we call *actuals*, are taken from the latest available vintage, i.e., the data released in 2012:Q1.

4. Results

In this section we report the relative accuracy of point and density forecasts, which is evaluated with the root mean squared forecast error (RMSFE) and the average log-predictive scores (LPSs), respectively.

⁴The results for the rolling forecasting scheme, which are available upon request, lead to broadly the same conclusions as those presented in section 4.

4.1 *Point Forecasts*

We begin our comparison by analyzing point forecasts. Table 1 reports the values for the RMSFEs, both using real-time and latest available data. As discussed above, the former case allows us to capture both sample and vintage effects, while the latter distills the sample effect. The numbers for the update 1Q scheme (our baseline) represent the values of the RMSFEs, while the remaining numbers are expressed as ratios so that values above (below) unity indicate that a given scheme underperforms (overperforms) the baseline. Moreover, to provide a rough gauge of whether the RMSFE ratios are significantly different from unity, we report the results of the Diebold-Mariano test.

The RTD results show that the accuracy of update 2Q and update 1Y schemes are not significantly different from the baseline. For the update 2Y variant the ratios for output, inflation, and the interest rate tend to be above unity, and in many cases significantly so. This brings us to the first conclusion: the parameters of the canonical medium-sized DSGE model can be reestimated once a year without a significant loss of point forecasts' accuracy for the main macro-variables. However, less frequent reestimation leads to a significant deterioration in the quality of predictions for these variables.

As regards the fixed scheme, the loss in forecast precision is very large. For instance, the RMSFEs for the interest rate go up by as much as 30 percent compared with the baseline. The ratios are even larger for output and the price level, exceeding 1.5 and 1.8 for the three-year horizon. Hence, our second conclusion is that evaluating the forecasting performance of DSGE models without updating their parameters can give a distorted picture.

The comparison of the RTD and LAD results for the baseline scheme indicates that the use of RTD generates the RMSFEs for output and inflation that are about 10 percent higher than in the LAD case. For the remaining updating schemes, except for the fixed one, the RMSFE ratios are very similar. Since there is no vintage effect for LAD, our third finding is that conclusion one (reestimate the model at least once a year) is driven mainly by the sample effect.

Table 1. Root Mean Squared Forecast Errors (RMSFEs)

	<i>H</i> = 1	<i>H</i> = 2	<i>H</i> = 4	<i>H</i> = 6	<i>H</i> = 8	<i>H</i> = 12
Real-Time Data Results						
<i>Output</i>						
Update 1Q	0.66	1.17	2.02	2.61	3.08	4.14
Update 2Q	1.00	1.00	1.00	1.00	1.00	1.01*
Update 1Y	1.01	1.01	1.01	1.01	1.02	1.01
Update 2Y	1.03	1.02	1.03*	1.04**	1.04***	1.04**
Fixed	1.01	0.99	1.14	1.35	1.51**	1.55**
<i>Prices</i>						
Update 1Q	0.26	0.47	0.87	1.30	1.77	2.77
Update 2Q	1.00	0.99	0.99	0.99	1.00	1.01
Update 1Y	1.00	0.99	0.99	1.00	1.00	1.02
Update 2Y	1.00	1.00	1.01	1.02	1.04	1.07**
Fixed	1.01	1.10	1.38**	1.61***	1.73***	1.81***
<i>Interest Rate</i>						
Update 1Q	0.12	0.22	0.38	0.46	0.52	0.57
Update 2Q	1.00	1.00	1.00	1.00	1.00	1.00
Update 1Y	1.00*	1.00	1.00	1.00	1.00	0.99
Update 2Y	1.02**	1.03**	1.02*	1.01	1.01	0.98
Fixed	1.27***	1.29***	1.30***	1.27***	1.22***	1.16***
Latest Available Data Results						
<i>Output</i>						
Update 1Q	0.61	1.07	1.92	2.50	2.88	3.65
Update 2Q	1.01	1.01	1.01	1.00	1.01	1.01
Update 1Y	1.01	1.02	1.01	1.01	1.02*	1.01
Update 2Y	1.03	1.03	1.03	1.03	1.04	1.03
Fixed	1.13*	1.11*	1.02	0.97	0.94	0.92
<i>Prices</i>						
Update 1Q	0.21	0.39	0.77	1.20	1.72	2.85
Update 2Q	1.01	1.01	1.01	1.00	0.99	0.99
Update 1Y	1.02	1.03	1.02	1.02	1.02	1.03
Update 2Y	1.03	1.05	1.05	1.04	1.04	1.05
Fixed	1.06	1.10	1.20**	1.31***	1.39***	1.48***

(continued)

Table 1. (Continued)

	$H = 1$	$H = 2$	$H = 4$	$H = 6$	$H = 8$	$H = 12$
Latest Available Data Results						
<i>Interest Rate</i>						
Update 1Q	0.12	0.23	0.39	0.48	0.54	0.59
Update 2Q	1.00	1.00	1.00	1.00	1.00	1.00
Update 1Y	1.02*	1.02	1.01	1.00	1.00	1.00
Update 2Y	1.01	1.02	1.02	1.02	1.01	0.99
Fixed	1.22**	1.23**	1.24***	1.26***	1.29***	1.35***
Notes: For the baseline scheme (Update 1Q) the RMSFEs are reported in levels, whereas for the remaining schemes they appear as the ratios so that the values above unity indicate that a given model has a higher RMSFE than the baseline. Asterisks ***, **, and * denote the 1, 5, and 10 percent significance levels of the two-tailed Diebold-Mariano test, where the long-run variance is calculated with the Newey-West method.						

4.2 Density Forecasts

We complement the discussion of point forecasts’ accuracy with an evaluation of density forecasts. The aim is to check to what extent the analyzed forecasts provide a realistic description of actual uncertainty. The quality of the density forecast is assessed with the average log-predictive score statistic, which for h -step-ahead forecasts from the j -th scheme is equal to the mean value of log of $p(y_{t+h}|t, j)$ computed with formula (3).

In tables 2 and 3 we report the values of average LPSs for each of the three key macroeconomic variables separately, and for their joint distribution. As before, we present the results for the models estimated with RTD as well as with LAD. The numbers for the baseline represent the levels, whereas the remaining numbers are expressed as differences so that values below zero indicate that a given scheme underperforms the baseline. To provide a rough gauge of whether these differences are significantly different from zero, we report the results of the Amisano and Giacomini (2007) test.

The RTD results for individual variables show that decreasing the frequency of model reestimation leads to a deterioration in the quality of density forecasts. However, the decrease in the fit, although in some cases statistically significant, is not large

Table 2. Average Univariate Log-Predictive Scores (LPSs)

	<i>H</i> = 1	<i>H</i> = 2	<i>H</i> = 4	<i>H</i> = 6	<i>H</i> = 8	<i>H</i> = 12
Real-Time Data						
<i>Output</i>						
Update 1Q	−1.121	−1.660	−2.197	−2.461	−2.628	−2.895
Update 2Q	−0.004	−0.002	−0.003	−0.003	−0.004	−0.008*
Update 1Y	−0.010***	−0.009	−0.012*	−0.013*	−0.012	−0.010
Update 2Y	−0.022***	−0.023**	−0.032***	−0.032***	−0.030**	−0.026*
Fixed	−0.123**	−0.099	−0.132	−0.231	−0.323*	−0.422**
<i>Prices</i>						
Update 1Q	−0.144	−0.797	−1.506	−1.933	−2.232	−2.635
Update 2Q	0.000	0.002	0.002	0.001	0.000	−0.004*
Update 1Y	−0.002	0.000	−0.001	−0.001	−0.003	−0.008
Update 2Y	−0.008*	−0.005	−0.008	−0.012*	−0.016**	−0.029***
Fixed	−0.134***	−0.204***	−0.297***	−0.355***	−0.392***	−0.439***
<i>Interest Rate</i>						
Update 1Q	0.324	−0.128	−0.549	−0.738	−0.841	−0.941
Update 2Q	−0.004**	−0.004**	−0.003	−0.003	−0.002	0.001
Update 1Y	−0.008***	−0.009***	−0.009***	−0.006	−0.003	0.003
Update 2Y	−0.019***	−0.023***	−0.024***	−0.021**	−0.018	0.006
Fixed	−0.175***	−0.202***	−0.211***	−0.194***	−0.162***	−0.108***
Latest Available Data						
<i>Output</i>						
Update 1Q	−1.082	−1.599	−2.148	−2.420	−2.574	−2.799
Update 2Q	−0.005	−0.007	−0.006	−0.003	−0.004	−0.004
Update 1Y	−0.010*	−0.013	−0.012	−0.012*	−0.015**	−0.011*
Update 2Y	−0.020*	−0.024	−0.028*	−0.030**	−0.035**	−0.028**
Fixed	−0.139***	−0.124***	−0.076	−0.055	−0.057	−0.042
<i>Prices</i>						
Update 1Q	0.013	−0.673	−1.431	−1.883	−2.202	−2.633
Update 2Q	−0.004	−0.004	−0.002	−0.001	0.000	0.000
Update 1Y	−0.009	−0.010	−0.010	−0.009	−0.010	−0.014
Update 2Y	−0.016	−0.020*	−0.023**	−0.023**	−0.023**	−0.028**
Fixed	−0.093***	−0.141***	−0.194***	−0.233***	−0.263***	−0.318***

(continued)

Table 2. (Continued)

	<i>H</i> = 1	<i>H</i> = 2	<i>H</i> = 4	<i>H</i> = 6	<i>H</i> = 8	<i>H</i> = 12
Latest Available Data						
Interest Rate						
Update 1Q	0.334	−0.122	−0.551	−0.742	−0.843	−0.933
Update 2Q	−0.004*	−0.004	−0.003	−0.001	0.000	0.002
Update 1Y	−0.012***	−0.014**	−0.009	−0.003	−0.001	0.001
Update 2Y	−0.015***	−0.019***	−0.021**	−0.020	−0.018	0.004
Fixed	−0.162***	−0.172***	−0.183***	−0.204***	−0.229***	−0.268***
Notes: For the baseline scheme (Update 1Q) LPSs are reported in levels, whereas for the remaining schemes they appear as the differences so that the values below zero indicate that a given model has a lower LPS than the baseline. Asterisks ***, **, and * denote the 1, 5, and 10 percent significance levels for the two-tailed Amisano and Giacomini (2007) test, where the long-run variance is calculated with the Newey-West method.						

Table 3. Average Multivariate Log-Predictive Scores (LPSs) for Output, Prices, and Interest

	<i>H</i> = 1	<i>H</i> = 2	<i>H</i> = 4	<i>H</i> = 6	<i>H</i> = 8	<i>H</i> = 12
Real-Time Data						
Update 1Q	−0.889	−2.503	−4.108	−4.960	−5.507	−6.246
Update 2Q	−0.007*	−0.004	−0.006	−0.007	−0.010*	−0.014**
Update 1Y	−0.020***	−0.018**	−0.025***	−0.026**	−0.025**	−0.015
Update 2Y	−0.050***	−0.052***	−0.070***	−0.073***	−0.068***	−0.044*
Fixed	−0.393***	−0.432***	−0.531***	−0.642***	−0.726***	−0.800***
Latest Available Data						
Update 1Q	−0.679	−2.300	−3.962	−4.847	−5.404	−6.122
Update 2Q	−0.012*	−0.014*	−0.011	−0.007	−0.007	−0.006
Update 1Y	−0.027**	−0.031**	−0.027**	−0.024*	−0.027**	−0.022
Update 2Y	−0.047***	−0.057***	−0.064***	−0.068**	−0.071**	−0.045*
Fixed	−0.370***	−0.390***	−0.366***	−0.353***	−0.362***	−0.392***
Notes: See notes to table 2.						

for the update 2Q and update 1Y schemes: the average value of the predictive likelihood is up to 1.3 percent lower than for the baseline, depending on the variable and horizon. For the multivariate

forecasts, which take into account the covariances, the differences in the average LPSs for the update 1Y scheme are somewhat higher, amounting to 2.6 percent, and statistically significant. For the update 2Y and fixed schemes, the LPS differences are significantly negative for most variables and horizons. Moreover, their values indicate a sizable decline in the average predictive likelihood, amounting in the update 2Y scheme to around 7 percent for the joint density of all three macro-variables. The loss under the fixed scheme is even larger, reaching 80 percent in the multivariate case. We interpret these results as confirming our conclusions formulated for point forecasts, i.e., that DSGE models should be reestimated at least once a year. Moreover, they also show that there are some gains from reestimating the model more frequently than once a year, especially if one is interested in the quality of multivariate density forecasts.

The comparison of the RTD and LAD results for the baseline scheme shows that the use of RTD decreases the accuracy of density forecasts for output and inflation, with the fall in the average LPSs ranging from 5 to 15 percent. The decline is most sizable for the short-term inflation forecasts. The comparison of the LPS differences for the remaining schemes indicates that they are broadly the same for the RTD and LAD cases. This confirms our earlier finding that the sample effect dominates the vintage effect.

4.3 Robustness Checks

The results presented above suggest that the parameter estimates of our benchmark model are quite stable. At least to some extent, this may be due to the prior assumptions borrowed from the original Smets and Wouters (2007) model, which are sometimes considered to be rather tight. To check if this is the case, we repeat our calculations with looser priors. More specifically, we increase the standard deviations of prior distributions for all structural parameters by 50 percent, except for the degree of indexation and capacity utilization cost curvature, for which the standard deviations are raised by 33 percent.⁵

The results of this robustness check are presented in table 4. They show that loosening the priors usually slightly improves the quality of short-term forecasts and quite sizably decreases it for longer

⁵For these parameters, increasing the standard deviations of the prior (beta) distributions by 50 percent would result in bimodality.

Table 4. Relative Forecast Accuracy of Model with Loose Priors

	<i>H</i> = 1	<i>H</i> = 2	<i>H</i> = 4	<i>H</i> = 6	<i>H</i> = 8	<i>H</i> = 12
Root Mean Squared Forecast Errors						
<i>Output</i>						
Update 1Q	0.64	1.12	1.95	2.59	3.21	4.65
Update 2Q	1.00	1.00	1.00	1.00	1.00	1.00
Update 1Y	1.01	1.00	1.01	1.01	1.00	1.00
Update 2Y	1.01	1.00	1.01	1.02*	1.02	1.02
Fixed	1.06	1.04	1.15	1.31	1.39*	1.34**
<i>Prices</i>						
Update 1Q	0.25	0.46	0.91	1.45	2.07	3.43
Update 2Q	1.00	1.00*	1.00	1.00	1.00	1.00
Update 1Y	1.00	1.00	1.00	0.99	0.99	0.99
Update 2Y	1.00	1.00	1.00	1.01	1.01	1.02
Fixed	1.09	1.26**	1.56***	1.71***	1.74***	1.71***
<i>Interest Rate</i>						
Update 1Q	0.11	0.21	0.36	0.46	0.52	0.60
Update 2Q	1.00	1.00	1.00	1.00	1.00	1.00
Update 1Y	1.01**	1.01*	1.01*	1.01*	1.01	1.00
Update 2Y	1.02*	1.02**	1.02**	1.02*	1.01	1.00
Fixed	1.41***	1.45***	1.43***	1.37***	1.30***	1.21***
Average Log-Predictive Scores						
<i>Output</i>						
Update 1Q	−1.098	−1.636	−2.172	−2.454	−2.652	−2.979
Update 2Q	−0.003	−0.001	−0.002	−0.003	−0.003	−0.005
Update 1Y	−0.008***	−0.007**	−0.007	−0.006	−0.005	−0.002
Update 2Y	−0.013***	−0.013**	−0.018***	−0.023**	−0.023*	−0.024*
Fixed	−0.139***	−0.120	−0.147	−0.218	−0.267**	−0.287**
<i>Prices</i>						
Update 1Q	−0.112	−0.755	−1.474	−1.930	−2.265	−2.731
Update 2Q	0.000	0.002	0.001	0.002	0.002	0.001
Update 1Y	−0.003	0.000	0.000	0.001	0.003	0.004
Update 2Y	−0.007	−0.002	−0.003	−0.004	−0.003	−0.006
Fixed	−0.184***	−0.279***	−0.391***	−0.447***	−0.471***	−0.493***

(continued)

Table 4. (Continued)

	<i>H</i> = 1	<i>H</i> = 2	<i>H</i> = 4	<i>H</i> = 6	<i>H</i> = 8	<i>H</i> = 12
Average Log-Predictive Scores						
<i>Interest Rate</i>						
Update 1Q	0.345	−0.107	−0.529	−0.724	−0.840	−0.972
Update 2Q	−0.004**	−0.005**	−0.004	−0.003	−0.002	0.002
Update 1Y	−0.011***	−0.012***	−0.012***	−0.011***	−0.008*	0.001
Update 2Y	−0.018***	−0.018***	−0.018***	−0.018***	−0.016**	−0.003
Fixed	−0.216***	−0.254***	−0.270***	−0.251***	−0.211***	−0.136***
<i>Three Variables</i>						
Update 1Q	−0.810	−2.404	−3.990	−4.854	−5.435	−6.256
Update 2Q	−0.008*	−0.004	−0.003	−0.004	−0.006	−0.005
Update 1Y	−0.020***	−0.014**	−0.012*	−0.012	−0.010	−0.004
Update 2Y	−0.035***	−0.027***	−0.027***	−0.031**	−0.032*	−0.035*
Fixed	−0.483***	−0.559***	−0.680***	−0.767***	−0.805***	−0.787***
Notes: See notes to tables 1 and 2.						

horizons. Looking at a comparison across alternative updating schemes, this variant suggests that reestimations can be carried out every two years without significantly affecting the quality of point forecasts for output and inflation. As regards the density forecasts, the main conclusions are broadly the same as those formulated using our baseline results.

Our next check is motivated by the absence of a financial sector in the benchmark Smets and Wouters (2007) model, even though one might argue that risk premium shocks can capture, at least to some extent, disruptions in financial intermediation. This simplified description of financial markets may significantly affect the forecasting performance, especially during times of financial stress. We address this concern by augmenting the benchmark model with the financial accelerator mechanism in the spirit of Bernanke, Gertler, and Gilchrist (1999). The exact implementation of this extension follows Del Negro, Giannoni, and Schorfheide (2014) and uses corporate bond spreads as an additional observable variable. The details can be found in the appendix.

Table 5 summarizes the results of this robustness check. In the short run, the quality of point forecasts for output and the interest

Table 5. Relative Forecast Accuracy of Model with Financial Frictions

	<i>H</i> = 1	<i>H</i> = 2	<i>H</i> = 4	<i>H</i> = 6	<i>H</i> = 8	<i>H</i> = 12
Root Mean Squared Forecast Errors						
<i>Output</i>						
Update 1Q	0.61	1.07	1.97	2.76	3.51	4.93
Update 2Q	1.01	1.00	1.01	1.01	1.01	1.01*
Update 1Y	1.02**	1.01**	1.02***	1.03***	1.02**	1.02**
Update 2Y	1.02**	1.02**	1.03***	1.03***	1.03**	1.03**
Fixed	1.22*	1.36*	1.47**	1.54**	1.57**	1.57***
<i>Prices</i>						
Update 1Q	0.26	0.52	1.20	2.12	3.25	5.86
Update 2Q	1.00	1.00	1.01	1.01	1.01	1.02
Update 1Y	1.01*	1.02*	1.03*	1.04*	1.05**	1.05**
Update 2Y	1.04***	1.06***	1.10***	1.11***	1.11***	1.10***
Fixed	1.21**	1.48***	1.91***	2.15***	2.28***	2.43***
<i>Interest Rate</i>						
Update 1Q	0.11	0.20	0.37	0.52	0.65	0.83
Update 2Q	1.00**	1.00	1.00	1.00	1.00	1.00
Update 1Y	1.00	1.00	1.01	1.02*	1.02*	1.02*
Update 2Y	1.01	1.02	1.02	1.02	1.02	1.03**
Fixed	1.32***	1.42***	1.58***	1.71***	1.81***	1.96***
Log-Predictive Scores						
<i>Output</i>						
Update 1Q	−1.071	−1.587	−2.139	−2.457	−2.682	−3.022
Update 2Q	−0.004	−0.001	−0.006*	−0.006	−0.006	−0.012*
Update 1Y	−0.009***	−0.010***	−0.020***	−0.021***	−0.018**	−0.021*
Update 2Y	−0.015***	−0.015***	−0.025***	−0.022**	−0.016	−0.023
Fixed	−0.138**	−0.203**	−0.308*	−0.425**	−0.557**	−0.877***
<i>Prices</i>						
Update 1Q	−0.157	−0.891	−1.699	−2.210	−2.591	−3.133
Update 2Q	−0.004	−0.003	−0.005	−0.006	−0.006	−0.007
Update 1Y	−0.022*	−0.025*	−0.030**	−0.032**	−0.031**	−0.031**
Update 2Y	−0.079**	−0.100**	−0.108**	−0.096***	−0.090***	−0.076***
Fixed	−0.335***	−0.446***	−0.613***	−0.723***	−0.801***	−0.921***

(continued)

Table 5. (Continued)

	<i>H</i> = 1	<i>H</i> = 2	<i>H</i> = 4	<i>H</i> = 6	<i>H</i> = 8	<i>H</i> = 12
Log-Predictive Scores						
<i>Interest Rate</i>						
Update 1Q	0.316	−0.127	−0.586	−0.852	−1.037	−1.271
Update 2Q	−0.002***	−0.002***	−0.002*	−0.002	−0.001	−0.002
Update 1Y	−0.006***	−0.007***	−0.010**	−0.009	−0.008	−0.009
Update 2Y	−0.018***	−0.021***	−0.021**	−0.017**	−0.016**	−0.024*
Fixed	−0.240***	−0.296***	−0.380***	−0.452***	−0.518***	−0.627***
<i>Three Variables</i>						
Update 1Q	−0.902	−2.633	−4.388	−5.346	−5.989	−6.897
Update 2Q	−0.009	−0.005	−0.017	−0.019*	−0.019*	−0.026*
Update 1Y	−0.033**	−0.030*	−0.046*	−0.054**	−0.056**	−0.059**
Update 2Y	−0.095***	−0.090*	−0.117*	−0.115**	−0.114**	−0.126***
Fixed	−0.548***	−0.634***	−0.835***	−1.062***	−1.313***	−1.899***
Notes: See notes to tables 1 and 2.						

rate improve. However, if we focus on longer horizons, the picture is reversed and the accuracy of predictions for prices clearly deteriorates across the board. Similar patterns emerge for density forecasts. These findings are in line with Del Negro and Schorfheide (2012) and Kolasa and Rubaszek (2015), who show that including financial frictions in the corporate sector usually generates less accurate forecasts, even though their quality improves during the Great Recession. As regards the frequency of reestimation in this alternative specification, the results suggest to do it twice rather than once a year if the main focus is on point forecasts for output and prices. Otherwise, and especially if the quality of density forecasts is the main target, the results are consistent with our baseline conclusions.

Finally, one might be concerned that our recommended interval between consecutive model reestimations might be driven by choosing the first quarter of the year in the 1Y and 2Y schemes as the time of parameter update, or by some special events occurring just before reestimation in the 1Y schedule (and hence one year before that done according to the 2Y scheme). The additional checks (not

reported but available upon request) prove that this is not the case and that our main conclusions are based on the intrinsic features of the alternative updating schemes.

5. Conclusions

The results of this study show that the common practice of policy-making institutions to reestimate DSGE models only occasionally is justified and does not lead to any sizable loss in forecast accuracy for the main macro variables as long as the interval between consecutive reestimations does not exceed one year. Such a frequency of updating the model parameters facilitates the communication between the modelers and policymakers and allows to save on model maintenance costs, without coming at the expense of forecast quality. The main source of deterioration in forecast precision when reestimations become less frequent is related to new data arrival (sample effect) rather than data revisions (vintage effect). Finally, according to our results, while assessing the forecasting performance of DSGE models, it might matter whether real-time or latest available data are used.

Appendix

List of Model Equations

This section lays out the full system of log-linearized equations that make up the DSGE model analyzed in the main text. The specification is the same as that used by Smets and Wouters (2007). The only difference concerns trend investment-specific technological progress, which we include to account for a secular trend in the real investment to output ratio observed in the data.

This modification means that for the model to have a well-defined steady state, the trending real variables need to be normalized both with neutral (γ) and investment-specific (γ_i) technological progress (see, e.g., Greenwood, Herzowitz, and Krusell 1997).

In particular, if we define $\gamma_y = \gamma\gamma_i^{\frac{\alpha}{1-\alpha}}$, the stationarization is as follows: $y_t = Y_t/\gamma_y^t$, $y_t^p = Y_t^p/\gamma_y^t$, $c_t = C_t/\gamma_y^t$, $i_t = I_t/(\gamma_y\gamma_i)^t$, $k_t = K_t/(\gamma_y\gamma_i)^t$, $k_t^s = K_t^s/(\gamma_y\gamma_i)^t$, $w_t = W_t/(P_t\gamma_y^t)$, $r_t^k = R_t^k\gamma_i^t/P_t$, and $q_t = Q_t\gamma_i^t$, where Y_t is output, Y_t^p is potential output, C_t is

consumption, I_t is investment, K_t is capital, K_t^s is capital services, W_t is nominal wage, R_t^k is the nominal rental rate on capital, Q_t is the real price of capital, and P_t is the price level. As regards the remaining endogenous variables showing up in the equations below, l_t stands for labor, r_t is the nominal interest rate, π_t is inflation, μ_t^p is price markup, μ_t^w is wage markup, and z_t is capital utilization.

The model is driven by seven stochastic disturbances: total factor productivity ε_t^a , investment-specific technology ε_t^i , risk premium ε_t^b , exogenous spending ε_t^g , price markup ε_t^p , wage markup ε_t^w , and monetary policy ε_t^r . The two markup shocks are assumed to follow ARMA(1,1) processes. All remaining disturbances are modeled as first-order autoregressions, except that the government spending shock additionally depends on the current innovation to total factor productivity. All variables presented in the equations below are expressed as log-deviations from the non-stochastic steady state. The parameters are defined in the “Calibration and Prior Assumptions” section of this appendix. Stars in subscripts indicate the steady-state values, which are functions of deep model parameters.

Aggregate resource constraint:

$$y_t = (1 - (\gamma_y \gamma_i - 1 + \delta) k_* y_*^{-1} - g_y) c_t + (\gamma_y \gamma_i - 1 + \delta) k_y i_t + r_*^k k_* y_*^{-1} z_t + g_y \varepsilon_t^g \quad (4)$$

Consumption Euler equation:

$$c_t = \frac{\lambda}{\lambda + \gamma_y} c_{t-1} + \frac{\gamma_y}{\lambda + \gamma_y} E_t c_{t+1} + \frac{\gamma_y (\sigma_c - 1) \frac{w_* l_*}{c_*}}{\sigma_c (\gamma_y + \lambda) (1 + \lambda_w)} (l_t - E_t l_{t+1}) - \frac{\gamma_y - \lambda}{\sigma_c (\gamma_y + \lambda)} (r_t - E_t \pi_{t+1}) + \varepsilon_t^b \quad (5)$$

Investment Euler equation:

$$i_t = \frac{1}{1 + \beta \gamma_y^{1-\sigma_c}} i_{t-1} + \frac{\beta \gamma_y^{1-\sigma_c}}{1 + \beta \gamma_y^{1-\sigma_c}} E_t i_{t+1} + \frac{1}{(1 + \beta \gamma_y^{1-\sigma_c}) \gamma_y^2 \gamma_i^2 \varphi} q_t + \varepsilon_t^i \quad (6)$$

Value of capital:

$$q_t = \frac{\beta(1-\delta)}{\gamma_y^{\sigma_c} \gamma_i} E_t q_{t+1} + \frac{\gamma_y^{\sigma_c} \gamma_i - \beta(1-\delta)}{\gamma_y^{\sigma_c} \gamma_i} E_t r_{t+1}^k - (r_t - E_t \pi_{t+1}) + \frac{\sigma_c(\gamma_y + \lambda)}{\gamma_y - \lambda} \varepsilon_t^b \quad (7)$$

Aggregate production function:

$$y_t = \phi_p(\alpha k_t^s + (1-\alpha)l_t + \varepsilon_t^a) \quad (8)$$

Capital services:

$$k_t^s = k_{t-1} + z_t \quad (9)$$

Optimal capacity utilization:

$$z_t = \frac{1-\psi}{\psi} r_t^k \quad (10)$$

Capital accumulation:

$$k_t = \frac{1-\delta}{\gamma_y \gamma_i} k_{t-1} + \frac{\gamma_y \gamma_i - 1 + \delta}{\gamma_y \gamma_i} i_t + (\gamma_y \gamma_i - 1 + \delta)(1 + \beta \gamma_y^{1-\sigma_c}) \gamma_y \gamma_i \varphi \varepsilon_t^i \quad (11)$$

Price markup:

$$\mu_t^p = \alpha(k_t^s - l_t) + \varepsilon_t^a - w_t \quad (12)$$

Phillips curve:

$$\pi_t = \frac{\iota_p}{1 + \beta \gamma_y^{1-\sigma_c} \iota_p} \pi_{t-1} + \frac{\beta \gamma_y^{1-\sigma_c}}{1 + \beta \gamma_y^{1-\sigma_c} \iota_p} E_t \pi_{t+1} - \frac{(1 - \beta \gamma_y^{1-\sigma_c} \xi_p)(1 - \xi_p)}{(1 + \beta \gamma_y^{1-\sigma_c} \iota_p) \xi_p ((\phi_p - 1) \varepsilon_p + 1)} \mu_t^p + \varepsilon_t^p \quad (13)$$

Input cost minimization:

$$r_t^k = -(k_t - l_t) + w_t \quad (14)$$

Wage markup:

$$\mu_t^w = w_t - \left(\sigma_l l_t + \frac{1}{\gamma_y - \lambda} (\gamma_y c_t - \lambda c_{t-1}) \right) \quad (15)$$

Real wage dynamics:

$$\begin{aligned} w_t = & \frac{1}{1 + \beta \gamma_y^{1-\sigma_c}} w_{t-1} + \frac{\beta \gamma_y^{1-\sigma_c}}{1 + \beta \gamma_y^{1-\sigma_c}} (E_t w_{t+1} - E_t \pi_{t+1}) \\ & - \frac{1 + \beta \gamma_y^{1-\sigma_c} \iota_w}{1 + \beta \gamma_y^{1-\sigma_c}} \pi_t + \frac{\iota_w}{1 + \beta \gamma_y^{1-\sigma_c}} \pi_{t-1} \\ & - \frac{(1 - \beta \gamma_y^{1-\sigma_c} \xi_w)(1 - \xi_w)}{(1 + \beta \gamma_y^{1-\sigma_c}) \xi_w ((\phi_w - 1) \varepsilon_w + 1)} \mu_t^w + \varepsilon_t^w \end{aligned} \quad (16)$$

Taylor rule:

$$r_t = \rho r_{t-1} + (1 - \rho) [r_\pi \pi_t + r_y (y_t - y_t^p)] + r_{\Delta y} (\Delta y_t - \Delta y_t^p) + \varepsilon_t^r \quad (17)$$

Data

The source of all data used to estimate the model is the Real-Time Data Set for Macroeconomists (RTDSM) database maintained by the Federal Reserve Bank of Philadelphia. The only exception is the short-term interest rate, which is not subject to revisions and is taken from the Federal Reserve Board statistics. The exact definitions follow below (RTDSM codes in parentheses).

- *Output*: Real gross domestic product (ROUTPUT) divided by civilian non-institutional population (POP).
- *Consumption*: Real personal consumption expenditures (RCON) divided by civilian non-institutional population (POP).
- *Investment*: Real gross private domestic non-residential investment (RINVBF) divided by civilian non-institutional population (POP).
- *Hours*: Aggregate weekly hours (H) divided by civilian non-institutional population (POP), normalized to average one over the estimation sample.

Table 6. Calibrated Parameters

Parameter	Value	Description
δ	0.025	Depreciation Rate
g_y	0.18	Exogenous Spending Share in Output
λ_w	1.5	Steady-State Wage Markup
ε_p	10	Kimball Aggregator Curvature in the Goods Market
ε_w	10	Kimball Aggregator Curvature in the Labor Market

- *Wages*: Nominal wage and salary disbursements (WSD) divided by civilian non-institutional population (POP) and deflated by the price index for gross domestic product (P).
- *Price Level*: Price index for gross domestic product (P).
- *Interest Rate*: Federal funds rate.

Measurement Equations

The following equations relate the model variables to their empirical counterparts defined in the “Data” section of this appendix:

$$\Delta \log Output_t = \gamma_y - 1 + y_t - y_{t-1}$$

(18)

$$\Delta \log Consumption_t = \gamma_y - 1 + c_t - c_{t-1}$$

(19)

$$\Delta \log Investment_t = \gamma_y \gamma_i - 1 + i_t - i_{t-1}$$

(20)

$$\log Hours_t = \bar{l} + l_t$$

(21)

$$\Delta \log Wages_t = \gamma_y - 1 + w_t - w_{t-1}$$

(22)

$$\Delta \log PriceLevel_t = \bar{\pi} + \pi_t$$

(23)

$$InterestRate_t = \beta^{-1} \gamma_y^{\sigma_c} (1 + \bar{\pi}) - 1 + r_t$$

(24)

Calibration and Prior Assumptions

The calibrated parameters are reported in table 6, while tables 7 and 8 describe the prior assumptions used in Bayesian estimation.

Table 7. Prior Assumptions—Structural Parameters

Parameter	Type	Mean	Std.	Description
φ	Normal	4.0	1.5	Investment Adjustment Cost Curvature
σ_c	Normal	1.5	0.37	Inv. Elasticity of Intertemporal Substitution
h	Beta	0.7	0.1	Habit Persistence
ξ_w	Beta	0.5	0.1	Calvo Probability for Wages
σ_l	Normal	2.0	0.75	Inv. Frisch Elasticity of Labor Supply
ξ_p	Beta	0.5	0.1	Calvo Probability for Prices
l_w	Beta	0.5	0.15	Wage Indexation
l_p	Beta	0.5	0.15	Price Indexation
ψ	Beta	0.5	0.15	Capacity Utilization Cost Curvature
ϕ	Normal	1.25	0.12	Steady-State Price Markup
r_π	Normal	1.5	0.25	Weight on Inflation in Taylor Rule
ρ	Beta	0.75	0.1	Interest Rate Smoothing
r_y	Normal	0.12	0.05	Weight on Output Gap in Taylor Rule
$r_{\Delta y}$	Normal	0.12	0.05	Weight on Output-Gap Change in Taylor Rule
$\bar{\pi}$	Gamma	0.62	0.1	Steady-State Inflation Rate
$100(\beta^{-1} - 1)$	Gamma	0.25	0.1	Rate of Time Preference
$\bar{l}$	Normal	0.0	2.0	Steady-State Hours Worked
$100(\gamma_y - 1)$	Normal	0.4	0.1	Trend Growth of Output
$100(\gamma_i - 1)$	Normal	0.3	0.1	Trend Growth of Relative Price of Investment
α	Normal	0.3	0.05	Capital Share

Table 8. Prior Assumptions—Shocks

Parameter	Type	Mean	Std.	Description
σ_a	Inv. Gamma	0.1	2.0	Volatility of Productivity Shock
σ_b	Inv. Gamma	0.1	2.0	Volatility of Risk Premium Shock
σ_g	Inv. Gamma	0.1	2.0	Volatility of Exogenous Spending Shock
σ_i	Inv. Gamma	0.1	2.0	Volatility of Investment-Specific Shock
σ_r	Inv. Gamma	0.1	2.0	Volatility of Monetary Policy Shock
σ_p	Inv. Gamma	0.1	2.0	Volatility of Price Markup Shock
σ_w	Inv. Gamma	0.1	2.0	Volatility of Wage Markup Shock
ρ_a	Beta	0.5	0.2	Persistence of Productivity Shock
ρ_b	Beta	0.5	0.2	Persistence of Risk Premium Shock
ρ_g	Beta	0.5	0.2	Persistence of Exogenous Spending Shock
ρ_i	Beta	0.5	0.2	Persistence of Investment-Specific Shock
ρ_r	Beta	0.5	0.2	Persistence of Monetary Policy Shock
ρ_p	Beta	0.5	0.2	Persistence of Price Markup Shock
ρ_w	Beta	0.5	0.2	Persistence of Wage Markup Shock
μ_p	Beta	0.5	0.2	Moving-Average Term in Price Markup
μ_w	Beta	0.5	0.2	Moving-Average Term in Wage Markup
ρ_{ga}	Beta	0.5	0.2	Impact of Productivity on Exogenous Spending

Introducing Financial Frictions

As one of the robustness checks considered in the paper, we augment the benchmark Smets and Wouters (2007) model with financial frictions as in Bernanke, Gertler, and Gilchrist (1999). They introduce an additional type of agents, called entrepreneurs, who manage capital, finance their operations by borrowing from banks owned by households, and are subject to idiosyncratic risk that can be observed by lenders only at a cost. This contracting friction results in an endogenous and time-varying wedge between the rate of return on capital and the risk-free rate.

Our exact specification follows Del Negro, Giannoni, and Schorfheide (2014) and is implemented by replacing the value of capital equation (7) with

$$E_t \{r_{t+1}^e - r_t\} = \zeta_{pr,b}(q_t + k_t - n_t) - \frac{\sigma_c(\gamma_y + \lambda)}{\gamma_y - \lambda} \varepsilon_t^b + \varepsilon_{\sigma,t}, \quad (25)$$

where $\varepsilon_{\sigma,t}$ is a shock to the standard deviation of idiosyncratic risk faced by entrepreneurs, r_t^e is the rate of return on capital defined as

$$r_t^e = \frac{r_*^k}{r_*^k + 1 - \delta} r_t^k + \frac{1 - \delta}{r_*^k + 1 - \delta} q_t - q_{t-1} + \pi_t, \quad (26)$$

and the law of motion for entrepreneurs' net worth n_t can be written as

$$\begin{aligned} n_t = & \zeta_{n,r^e}(r_t^e - \pi_t) - \zeta_{n,r}(r_{t-1} - \pi_t) + \zeta_{n,qk}(q_{t-1} + k_{t-1}) \\ & + \zeta_{n,n}n_{t-1} - \frac{\zeta_{n,\sigma_\omega}}{\zeta_{pr,\sigma_\omega}} \varepsilon_{\sigma,t-1}. \end{aligned} \quad (27)$$

The family of parameters ζ showing up in the new equations depends on the deep parameters of the augmented model, including the following four describing the financial frictions block: the survival rate of entrepreneurs ν , steady-state standard deviation of idiosyncratic productivity, monitoring costs, and transfers from households to entrepreneurs. Conditional on other model parameters, the last three of the four new parameters can be uniquely pinned down by the debt elasticity of the external finance premium $\zeta_{pr,b}$ and two steady-state proportions, i.e., the quarterly bankruptcy rate F_* and

Table 9. Prior Assumptions—Financial Frictions Block

Parameter	Type	Mean	Std.	Description
$r_*^e - r_*$	Gamma	0.50	0.025	Steady-State External Finance Premium
$\zeta_{pr,b}$	Beta	0.05	0.005	Premium Elasticity with Regard to Debt
σ_σ	Inv. Gamma	0.10	2.000	Volatility of Risk Shock
ρ_σ	Beta	0.50	0.200	Persistence of Risk Shock

the premium $r_*^e - r_*$.⁶ As in Del Negro, Giannoni, and Schorfheide (2014), we fix F_* at 0.008, ν at 0.99, and estimate $\zeta_{pr,b}$, $r_*^e - r_*$, as well as the standard deviation and autocorrelation of $\varepsilon_{\sigma,t}$. Our prior assumptions are summarized in table 9. They are taken from Del Negro, Giannoni, and Schorfheide (2014), except for the risk-shock characteristics, which we parameterize identically to other AR(1) shocks in the benchmark model.

Finally, while estimating the augmented model, we additionally use the time series on credit spreads, defined as the difference between the Moody’s seasoned Baa corporate bond yields and the ten-year Treasury note yields (source: Federal Reserve Board statistics) and linked to the model concept of the external finance premium via the following measurement equation:

$$Spread_t = r_*^e - r_* + E_t \{ r_{t+1}^e - r_t \}.$$

(28)

References

Adolfson, M., J. Lindé, and M. Villani. 2007. “Forecasting Performance of an Open Economy DSGE Model.” *Econometric Reviews* 26 (2–4): 289–328.

Amisano, G., and R. Giacomini. 2007. “Comparing Density Forecasts via Weighted Likelihood Ratio Tests.” *Journal of Business and Economic Statistics* 25: 177–90.

⁶See the technical appendix to Del Negro and Schorfheide (2012) for derivations.

- Bernanke, B. S., M. Gertler, and S. Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework." In *Handbook of Macroeconomics*, Vol. 1C, ed. J. B. Taylor and M. Woodford, 1341–93 (chapter 21). Elsevier.
- Christoffel, K., G. Coenen, and A. Warne. 2010. "Forecasting with DSGE Models." ECB Working Paper No. 1185.
- Croushore, D., and T. Stark. 2001. "A Real-Time Data Set for Macroeconomists." *Journal of Econometrics* 105 (1): 111–30.
- Del Negro, M., M. Giannoni, and F. Schorfheide. 2014. "Inflation in the Great Recession and New Keynesian Models." NBER Working Paper No. 20055.
- Del Negro, M., and F. Schorfheide. 2012. "DSGE Model-Based Forecasting." In *Handbook of Economic Forecasting*, Vol. 2A, ed. G. Elliott and A. Timmermann, 57–140 (chapter 2). Elsevier.
- Edge, R. M., M. T. Kiley, and J.-P. Laforte. 2010. "A Comparison of Forecast Performance between Federal Reserve Staff Forecasts, Simple Reduced-Form Models, and a DSGE Model." *Journal of Applied Econometrics* 25 (4): 720–54.
- Geweke, J., and G. Amisano. 2014. "Analysis of Variance for Bayesian Inference." *Econometric Reviews* 33 (1–4): 270–88.
- Giannone, D., F. Monti, and L. Reichlin. 2010. "Incorporating Conjunctural Analysis in Structural Models." In *The Science and Practice of Monetary Policy Today*, ed. V. Wieland, 41–57. Springer-Verlag.
- Greenwood, J., Z. Hercowitz, and P. Krusell. 1997. "Long-Run Implications of Investment-Specific Technological Change." *American Economic Review* 87 (3): 342–62.
- Kolasa, M., and M. Rubaszek. 2015. "Forecasting Using DSGE models with Financial Frictions." *International Journal of Forecasting* 31 (1): 1–19.
- Kolasa, M., M. Rubaszek, and P. Skrzypczynski. 2012. "Putting the New Keynesian DSGE Model to the Real-Time Forecasting Test." *Journal of Money, Credit and Banking* 44 (7): 1301–24.
- Rubaszek, M., and P. Skrzypczynski. 2008. "On the Forecasting Performance of a Small-Scale DSGE Model." *International Journal of Forecasting* 24 (3): 498–512.
- Smets, F., and R. Wouters. 2003. "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association* 1 (5): 1123–75.

- . 2007. “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach.” *American Economic Review* 97 (3): 586–606.
- Swamy, P. A. V. B., and G. J. Schinasi. 1986. “Should Fixed Coefficients Be Reestimated Every Period for Extrapolation?” International Finance Discussion Paper No. 287, Board of Governors of the Federal Reserve System.
- Wieland, V., and M. H. Wolters. 2012. “Forecasting and Policy Making.” In *Handbook of Economic Forecasting*, Vol. 2, ed. G. Elliott and A. Timmermann, 239–325 (chapter 5). Elsevier.