Volume 5, Number 2 June 2009

IJCB INTERNATIONAL JOURNAL OF CENTRAL BANKING Volume 5, Number 2 June 2009

IJCB

INTERNATIONAL JOURNAL OF CENTRAL BANKING

International Journal of Central Banking

Volume 5, Number 2 June 2009

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

The *International Journal of Central Banking* is published quarterly (ISSN: 1815-4654). Online access to the publication is available free of charge at **www.ijcb.org**. Individual print subscriptions are available at an annual rate of $100 (USD).

Print subscription orders may be placed online at www.ijcb.org, by phone (+49 69 1344 7623), via fax (+49 69 1344 8553), 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
DG Research
European Central Bank
Postfach 16 03 19
D-60066 Frankfurt
Germany

Phone: +49 69 1344 7623
Fax: +49 69 1344 8553
E-mail: editor@ijcb.org

ISSN: 1815-4654

Modeling Bank Senior Unsecured Ratings: A Reasoned Structured Approach to Bank Credit Assessment*

Spyros Pagratis[a] and Marco Stringa[b]
[a]Athens University of Economics and Business
[b]Bank of England

This paper studies the impact of bank-specific financial indicators and macroeconomic variables on bank senior unsecured ratings by Moody's. Controlling for bank financial characteristics, we find significant evidence of procyclicality in bank ratings stemming from lagged interaction effects between the real output gap and the credit gap. In particular, macroeconomic slowdowns that follow credit booms tend to imply lower ratings. Similarly, when credit expansion above a trend is followed by strong economic performance, bank ratings tend to increase. Bank ratings also appear to correlate positively with the slope of the yield curve and tend to increase with sovereign ratings, market share of lending, and bank size. Given the ongoing debate on the importance, timeliness, and information content of credit ratings in general—and those assigned to banks in particular—the paper addresses a topic that is of great importance to central banks, regulators, and risk managers.

JEL Codes: G21, G24, C25.

*We wish to thank Viral Acharya, Nicola Anderson, Andy Haldane, Gikas Hardouvelis, Nada Mora, Silvia Pezzini, Hyun Song Shin, Elias Tzavalis, an anonymous referee, and seminar participants at the Bank of England and EFMA 2008 for insightful comments. The views expressed in this paper are those of the authors and are not necessarily endorsed by the Bank of England. Any errors also remain our responsibility. Corresponding author: Spyros Pagratis, Athens University of Economics and Business, 76 Patission Street, 10434 Athens, Greece; E-mail: spagratis@aueb.gr.

1. Introduction

Bank credit ratings are metrics of bank creditworthiness that market participants focus on. Among other types of ratings (e.g., ratings of short-term debt, subordinated debt, or bank financial strength ratings), senior unsecured ratings attract particular attention due to their wide use in financial covenants and regulatory rules, defining portfolio allocation mandates for asset managers, triggers in securitization transactions, and the risk weighting for assets under the standardized approach of Basel II. Changes in those ratings are likely to shape market sentiment, and adverse changes may exacerbate any difficulties faced by banks during periods of stress. Bank senior unsecured ratings could come to a sharper focus in periods when banks are facing challenging financial conditions. Such periods may coincide with heightened macroeconomic uncertainty, as market participants may reassess financial risks and possibly retrench from certain types of financing.

Recent downgrades of banks were attributed by rating agencies to more difficult economic conditions, in conjunction with a deterioration in asset quality and reduction in the fair value of certain types of financial instruments. From a policy perspective, procyclicality in bank ratings would be an undesirable outcome.[1] At the very least, a deterioration in bank ratings that coincides with an economic downturn might impart a blow to confidence in the sector at a time when it is vulnerable to negative sentiment due to heightened macroeconomic uncertainty. That could eventually have implications for the real economy if tightening lending criteria to banks, in response to lower ratings, would exacerbate tight credit conditions in the economy.

Analysis on how bank senior unsecured ratings respond to bank-specific financial indicators and macroeconomic developments would be useful both to market participants and to policymakers. Considering in particular how ratings respond to fluctuations in macroeconomic and financial conditions could add to the debate about factors that may amplify fluctuations in the credit cycle and the

[1] In line with Amato and Furfine (2004), ratings are considered *procyclical* if they tend to be higher when the economy expands and lower during economic downturns, after controlling for firm-specific factors.

real economy. This paper offers a positive economic analysis of factors that may have an impact on bank senior unsecured ratings, avoiding normative implications of the rating process. Following Amato and Furfine (2004) and Blume, Lim, and MacKinlay (1998), we use an ordered probit framework to predict future ratings based on currently available information.

Explanatory variables for the empirical model include key financial indicators from banks' published accounts and measures of domestic economic activity and general financial conditions. As a modeling choice, we focus on the rating approach by one rating agency only (i.e., Moody's) given our familiarity with their rating policies and the availability of sufficient data to estimate the empirical model. Consequently, our results may not necessarily hold for other rating agencies and do not offer a judgment on Moody's rating performance relative to other agencies.

We take an informed approach to Moody's credit assessment of banks, both in terms of variable selection and model specification. The implicit assumption we make is that a single rating methodology is applied consistently across banks, as described in a number of public documents by Moody's (see, e.g., Moody's Investors Service 2002a, 2002b, 2006a, 2006b). The preferred model is estimated using a panel of annual data for 293 banks from thirty-three countries, covering the period 1999–2006. In employing bank-specific financial indicators to predict bank ratings, special attention is given to differences in accounting standards across jurisdictions and their possible impact on reported figures by banks. In particular, we control for country effects and we distinguish between banks that report their financial statements under International Financial Reporting Standards (IFRS) or national Generally Accepted Accounting Practices (GAAP).

However, accounting-based financial indicators are likely to be subject to cyclical variations. Albertazzi and Gambacorta (2006), for example, find evidence of procyclicality in measures of profitability and asset quality. Hence, the extent to which bank-specific financial indicators are informative about future ratings could depend on the phase of the business cycle. Higher profits, for example, and lower levels of write-offs during benign economic conditions are often supported by high business volumes, ample availability of credit, and strong asset valuations. But they could also mask vulnerabilities building up on banks' balance sheets, which could crystallize if economic conditions deteriorate. Therefore, increased profitability and perceived high asset quality during good times may not necessarily imply higher ratings if ratings take a long-term perspective, i.e., if they look through the cycle. In contrast, banks that are able to perform better relative to their peers regardless of the phase of the business cycle, should, in theory, be those that attract the higher ratings. In order to control for the phase of the business cycle, we consider the deviations (gaps) of domestic real GDP from a trend. We also consider the term spread (slope) of the yield curve as a forward-looking indicator of domestic economic activity, as suggested by Estrella and Hardouvelis (1991).

In addition, we consider country-level measures of financial imbalances that have been identified in the literature as forward-looking indicators of banking-sector vulnerabilities. Following Borio and Lowe (2002), variables that could help improve the explanatory power of the empirical model include measures of domestic credit expansion by the banking sector, asset market valuations, and foreign exchange mismatches. Deviations of those variables from a trend (gaps) are used to capture the potential buildup of financial imbalances. Borio and Lowe suggest that credit expansion above normal levels could *sow the seeds* of a subsequent deterioration in banks' risk profile (and possibly ratings) if economic conditions deteriorate. This view is supported by a number of studies (see, e.g., Dell' Ariccia, Igan, and Laeven 2008; Jiménez and Saurina 2006; Lown, Morgan, and Rohatgi 2000) showing that excessive credit expansions tend to be associated with relaxations in lending standards. Loose credit policies could lower banks' asset quality, which could then lead to credit risk crystallizing on banks' balance sheets as the economy enters a period of slowdown.

By considering cyclical variations in economic activity and financial conditions, we are able to examine whether senior unsecured ratings assigned to banks by Moody's look through the cycle. This adds to the empirical literature that investigates whether corporate ratings, more broadly, tend to be procyclical (see, e.g., Amato and Furfine 2004; Cantor and Mann 2003). We consider two possible channels through which procyclicality in bank ratings could manifest itself: (i) lending boom-bust episodes and (ii) cyclical fluctuations in bank-specific measures of profitability and asset quality.

In theory, bank ratings that look through those channels would also tend to look through the cycle. Otherwise, any inherent procyclicality in the financial system, as well as in measures of bank profitability and asset quality, could also translate into procyclical bank ratings.

Controlling for bank financial characteristics, we find empirical support for procyclicality in bank ratings stemming from lead-lag effects between the credit cycle and the business cycle. In particular, macroeconomic slowdowns (negative real output gaps, in one-year lag) that follow credit booms (positive credit gaps, in two-year lag) tend to imply lower ratings. Similarly, when credit expansion above a trend is followed by strong economic performance, bank ratings tend to be higher. We also find significant evidence that bank ratings internalize cyclical variations in asset quality by penalizing low asset quality more aggressively in good times than in periods of economic slowdown. However, no significant evidence is found for a similar *filter* applied to bank profitability, with ratings showing similar sensitivity to shocks in earnings both during economic booms and in periods of economic slowdown. Finally, bank ratings appear to correlate positively with the slope of the yield curve and tend to increase with sovereign ratings, market share of lending, and bank size. Overall, the estimated coefficients of bank-specific financial indicators are statistically significant and consistent with economic intuition and public statements by Moody's. The model performs well both in and out of sample and the results are robust to alternative model specifications.

The structure of the paper is as follows. Section 2 provides an outline of Moody's approach in assigning bank ratings. Section 3 describes the data and defines explanatory variables for the empirical model. Section 4 outlines the ordered probit methodology and discusses issues relating to model and sample selection. Section 5 presents the results and discusses robustness checks. Section 6 concludes.

2. Rating Methodology

Moody's aims for globally consistent rating scales, providing a rank ordering of risks associated with the ability and willingness of borrowers to meet debt obligations in full and on a timely basis. Moody's produces bank ratings on the basis of a general-to-specific approach (Moody's Investors Service 1999). Firstly, they examine the economic environment of the country of domicile and they consider strengths and weaknesses of the industry as a whole. Then, they consider debtor-specific characteristics in relation to peer groups. As part of their credit assessment process, Moody's has access to nonpublic information, either under the U.S. Regulation Fair Disclosure, which prohibits selective disclosure of nonpublic information but provides a conditional exception for rating agencies, or through private confidentiality agreements with issuers.[2]

Bank ratings by Moody's are based on five main areas of fundamental analysis: capital adequacy, asset quality, management, earnings and profitability, and funding and liquidity (CAMEL). Capital is aimed to absorb unexpected losses. After profitability, capital provides the second buffer to banks to withstand financial shocks, and the higher these buffers, the higher the resilience of banks to shocks. Asset quality is central to bank solvency and is therefore important for maintaining confidence among investors. Management quality, the most challenging category to capture quantitatively, spans a wide range of qualitative characteristics, such as cost efficiency, experience, and integrity—all of which affect the bank's riskiness and quality of earnings. Earnings capacity relates to the franchise value and profitability of the bank. It offers a first line of defense to debtholders in periods of stress and is considered by Moody's to be the *cornerstone of bank credit assessment* (see, e.g., Moody's Investors Service 2002a). Liquidity is relevant to bank credit assessment because banks are susceptible to customers' loss of confidence and sudden withdrawals of funds. Because of banks' maturity transformation role, high leverage, and intrinsic opaqueness, liquidity problems may become funding problems and even lead to insolvency.[3]

Moody's also aims to produce ratings that accommodate at the same time both rating stability and prudence, which are two widely

[2]Fight (2001) reports that more than 90 percent of rated firms reveal nonpublic information to the rating agencies.

[3]Flannery, Kwan, and Nimalendran (2004) and Morgan (2002) provide evidence on the opaqueness of banking institutions.

recognized criteria for rating.[4] In addition, senior unsecured ratings incorporate the probability and expected scale of official safety nets.[5] That could lead to more stable bank ratings over time—relative to what is predicted by an empirical model that is based on financial indicators and macroeconomic variables, as those ratings may react only partially, and sometimes not at all, to standard measures of bank financial health.

Based on these observations about the rating process, we now turn to discuss data and variable definitions for the empirical model.

3. Data and Variable Definitions

In order to estimate the empirical model presented in section 4, we employ a panel of annual data for 293 banks from thirty-three countries, covering the period 1999–2006. We consider senior unsecured ratings that are assigned to banks by Moody's as of the end of each calendar year. Data on bank ratings are obtained from Moody's Investors Service, spanning the rating spectrum of seventeen categories from Aaa through Caa3 in the familiar Moody's symbol system.[6] Because estimation of the empirical model requires a sufficient number of observations per rating class, we group banks into ten rating categories, where we focus our analysis. We assign the value 10 if a bank has a rating of Aaa–Aa1; 9 if Aa2; 8 if Aa3; 7 if A1; 6 if A2; 5 if A3; 4 if Baa1; 3 if Baa2–Baa3; 2 if Ba1–Ba3; and 1 if B1 or below.

Bank-specific financial ratios and country-level financial and macroeconomic indicators are used to define explanatory variables. We also consider sovereign ratings to control for sovereign credit

[4]Cantor and Mann (2007) argue that rating stability is desirable because rating changes can lead to actions by investors that are costly to reverse, primarily due to rating-based triggers in loan covenants and portfolio restrictions. Prudence is intrinsic to the interests of debtholders, which rating agencies aim to represent, implying that agencies prefer to err on the conservative side.

[5]By official safety nets we mean bank regulation and supervision, as well as emergency liquidity assistance by the official sector if a bank is in financial distress.

[6]Following Amato and Furfine (2004), we focus on rating actions that involve some degree of judgment by Moody's. Therefore, we eliminate observations of banks in the state of default, given that default is defined mechanically on the basis of well-known criteria. Nevertheless, bank defaults in the sample period that we examine are extremely rare events.

risk in a bank's country of domicile. Data on bank financial ratios are obtained from Bankscope.[7] Country-level data include finally revised figures from the IMF International Financial Statistics. Data on domestic interest rates are obtained from Global Financial Data. As with bank ratings, data on sovereign ratings are obtained from Moody's Investors Service. Next we discuss bank- and country-level explanatory variables.

3.1 Bank-Level Variables

Bank-specific variables are constructed using five key financial ratios (one for each CAMEL category): shareholders' equity/total assets; loan loss reserves/net interest income; operating costs/total assets; pre-tax, pre-provision profits/total assets; and deposits/customer loans. Using criteria that we discuss in section 4, we select these ratios from a set of financial indicators that we present in table 1. Table 1 also shows descriptive statistics on financial ratios to facilitate the discussion of results in section 5. In addition, we consider measures for market share and bank size, and we control for regional effects, such as country and sectoral concentrations. Regional effects are defined in terms of groups of countries that are shown in table 2, which also presents the regional distribution of observations and banks in the sample.

More specifically, the first bank-level indicator that we consider is capital. Managers target capital ratios that balance the requirements of many constituents, including shareholders, regulators, and rating agencies.[8] Therefore, instead of focusing on the rating impact of capital ratios per se, we consider percentage deviations of capital ratios from a *target.* The intuition is that capital adequacy is considered by Moody's in conjunction with the overall risk profile of a bank and the quality of its earnings (see, e.g., Moody's Investors Service 2002a, 2002b, 2006a). Hence, the impact of capital on ratings would depend on a bank's capital position relative to an appropriate target. Target ratios are estimated using a panel regression of capital

[7]Bankscope reports consolidated balance-sheet and income-statement information from banks' published accounts.

[8]The ability of banks to actively target a desired capital ratio has increased over recent years by the significant growth of structured credit products.

Table 1. List of Bank Financial Indicators Considered

Variable	Mean	Mean Abs. Deviation
Capital (%)		
Tier 1 Capital/Risk-Weighted Assets	12.48	6.60
Tier 1 and Tier 2 Capital/Risk-Weighted Assets	12.81	2.40
Shareholders' Equity/Total Assets	7.00	2.94
Shareholders' Equity/Loans	9.36	2.96
Shareholders' Equity/Total Liabilities	5.79	1.80
Asset Quality (%)		
Loan Loss Reserves/Loans + Loan Loss Reserves	1.88	0.92
Loan Loss Reserves/Net Interest Income	20.55	18.18
Loan Loss Reserves/Impaired Loans	152.78	107.57
Impaired Loans/(Loans + Loan Loss Reserves)	2.99	2.49
Loan Write-Offs/(Loans + Loan Loss Reserves)	1.08	1.27
Management (%)		
Operating Costs and Provisions/Total Assets	1.57	0.67
Operating Costs/Income Before Provisions	46.65	9.55
Operating Costs/Total Assets	2.19	1.07
Earnings (%)		
Net Interest Income/Total Earning Assets	1.31	0.63
Net Interest Income/Total Assets	1.21	0.56
Other Operating Income/Total Assets	0.53	0.39
Pre-Tax Profits/Total Assets	1.07	0.80
Pre-Tax, Pre-Provision Profits/Total Assets	1.53	0.86
Net Income/Total Assets	0.07	0.62
Net Income/Shareholders' Equity	2.04	8.96
Off-Balance-Sheet Exposures/Total Assets	21.43	21.81
Liquidity (%)		
Money Lent to Banks/Money Borrowed from Banks	87.44	46.94
Customer Loans/Total Assets	41.36	12.45
Customer Loans/Short-Term Liabilities	56.59	14.01
Liquid Assets/Short-Term Liabilities	9.00	4.79
Liquid Assets/Total Debt Exc. Capital Instruments	7.18	3.84

Note: The first column reports the list of bank-specific financial ratios that we consider as explanatory variables in the empirical model. The second and third columns report the sample mean and mean absolute deviation, respectively, for each financial ratio.

Table 2. Number of Banks and Bank-Year Observations by Country Group

Country Group	Number of Observations	Number of Banks
United States, Canada	301	57
Denmark, Finland, Norway, Sweden	82	17
United Kingdom	78	18
Ireland, Portugal, Spain	151	31
Netherlands	53	11
Belgium, France, Luxembourg	43	11
Austria, Switzerland	60	13
Cyprus, Greece	43	8
Germany	105	26
Italy	65	16
Australia	57	10
Japan	133	22
Indonesia, Korea, Malaysia, Thailand	64	19
China, Hong Kong, India, Kazakhstan, Philippines, Russia, Singapore	134	34
Total	**1,369**	**293**

ratios on a set of explanatory variables that control for differences in business mix, domestic economic conditions, and accounting policies. The estimation results for target ratios are presented in table 3.[9] We allow deviations of capital ratios from the estimated target to have an impact on ratings in a nonlinear fashion. Following Blume, Lim, and MacKinlay (1998), we model the relationship between ratings and the percentage deviation from the capital target as piecewise-linear. Let C_{it} be the percentage deviation of capital ratio from the

[9]Estimation results in table 3 show that, ceteris paribus, banks with higher net interest income relative to other operating income tend to have higher capital ratios. This is not surprising given that net interest income relies on capital-intensive assets (e.g., loans), while other operating income, such as trading income and fees and commissions, typically depends on less capital-intensive business. The estimation results also show that banks tend to hold higher capital ratios in an economic slowdown (i.e., when the GDP gap is negative). This effect is associated with a deleveraging process by banks in a downturn that is documented in a number of studies (see, e.g., Shin and Adrian 2007).

Table 3. Linear Regression Estimates for Target Capital Ratios, for the Period 1999–2006

DEPENDENT VARIABLE: *Shareholders' Equity/Total Assets*	Coefficient	Robust Std. Error	z-stat.
INDEPENDENT VARIABLES:			
Loan Loss Reserves/Total Assets	−2.044***	0.301	−6.80
Net Interest Income/Total Assets	1.674***	0.511	3.27
Other Operating Income/Total Assets	0.472***	0.109	4.31
Off-Balance-Sheet Exposures/Total Assets	0.004	0.004	0.84
IFRS Bank (Dummy): 1 if consolidated accounts in IFRS	−0.704*	0.372	−1.89
Domestic Economic Slowdown (Dummy): 1 if real GDP gap < 0	0.428*	0.251	1.71
Country Dummies			
Belgium, France, Luxembourg	−1.811***	0.677	−2.67
Germany	−1.316	1.037	−1.27
United Kingdom	−1.158*	0.660	−1.75
Australia	−1.085	0.664	−1.63
Austria, Switzerland	−0.944	0.846	−1.12
Italy	−0.793*	0.473	−1.68
Netherlands	−0.496	0.736	−0.67
Ireland, Portugal, Spain	−0.046	0.506	−0.09
Japan	0.064	0.814	0.08
Indonesia, Korea, Malaysia, Thailand	0.168	1.011	0.17
Cyprus, Greece	0.195	0.563	0.35
Denmark, Finland, Norway, Sweden	1.763	1.089	1.62
China, Hong Kong, India, Kazakhstan, Philippines, Russia, Singapore	2.484***	0.808	3.07
Constant	3.243**	1.415	2.29

Note: The model is estimated using a data panel of 1,369 observations, for the period 1999–2006. The data panel includes published-accounts data of 293 banks from thirty-three countries (grouped in fourteen regions) and macroeconomic information.

*Significant at 10 percent; **significant at 5 percent; ***significant at 1 percent. The first column reports the estimated coefficients of explanatory variables in the model. The second and third columns report robust standard errors and z-statistics.

estimated target for bank i in year t. We consider three new capital variables cj_{it} $(j = 1, 2, 3)$ such that

$$C_{it} = \sum_{j=1}^{3} cj_{it} \tag{1}$$

with cj_{it} defined as follows:

	$c1_{it}$	$c2_{it}$	$c3_{it}$
$C_{it} \in [0, +\infty)$	C_{it}	0	0
$C_{it} \in [-15, 0)$	0	C_{it}	0
$C_{it} \in (-\infty, -15)$	0	-15	$C_{it} + 15$.

By dividing C_{it} into three ranges, $(-\infty, -15\%)$, $[-15\%, 0)$, and $[0, +\infty)$, we allow deviations from the target to have a different marginal impact on ratings.[10] Moreover, we are able to examine whether large negative deviations from the target convey any additional information about ratings or reflect factors unrelated to ratings, such as model error from the panel regression. In the latter case, the coefficient for $c3_{it}$ in the empirical model would not be statistically different than zero.

As a measure of asset quality, we consider the ratio of loan loss reserves to net interest income (*LLR/NII*). The intuition is that net interest margins must appropriately remunerate for the risks undertaken by the bank. An increase in this ratio would imply that interest margins do not sufficiently compensate for risks in the loan book. Hence, the higher that ratio, the lower is asset quality. However, when net interest income (*NII*) is negative, the ratio of loan loss provisions to net interest income becomes meaningless. Hence, we consider the sign of interest income to define our asset-quality variable as follows:

Asset-Quality Variable	$(LLR/NII) \geq 0$	$(LLR/NII) < 0$
$NII \geq 0$	(LLR/NII)	(LLR/NII)
$NII < 0$	0	–

If *NII* is positive, then provisions and *LLR/NII* have the same sign and we set the asset-quality variable equal to *LLR/NII.* If *NII* is negative and *LLR/NII* is positive, then provisions are negative and the asset-quality variable is set equal to zero. There is only one bank-year observation corresponding to such an event. Finally, if both *NII* and *LLR/NII* are negative, then provisions are positive

[10]For example, holding more capital may not necessarily lead to higher ratings. But a weakly capitalized bank relative to a target may be downgraded or be forced to increase its capital base to avoid a downgrade.

and there is no remuneration for risks that the bank undertakes. This event corresponds to five bank-year observations, which are omitted from the sample.

Management quality is an area particularly difficult to measure quantitatively. A possibility is to proxy management quality by using measures of cost efficiency. Cost ratios have attracted the attention of analysts as banks seek to cut costs and improve their operational efficiency. In order to limit the possibility of colinearity problems with other ratios, such as asset quality and profitability variables, we consider the ratio of operating costs to total assets.

As a profitability variable, we employ the ratio of pre-tax, pre-provision profits to total assets. Pre-tax, pre-provision profits are Moody's favorite indicators of earnings-generating power (Moody's Investors Service 2002a, 2006a). By adding back provisions into profits, the profitability variable controls for the profit margin that is available to debtholders to absorb adverse shocks. It also has the further advantage of avoiding obvious colinearity problems between the profitability and asset-quality ratio.

As a composite measure of a bank's liquidity and funding position, we consider its *funding gap*. This is defined as the difference of (customer loans)–(short-term liabilities), expressed as a percentage of customer loans. We consider as short-term liabilities all financial liabilities with remaining maturity of less than one year. Assuming that customer loans are typically long-term and illiquid assets, the higher the funding gap, the more illiquid the bank would be and possibly more vulnerable to a classic bank run à la Diamond and Dybvig (1983). Nevertheless, a too-low funding gap could be associated with excess liquidity and inefficient employment of financial resources. In the long run, that could weaken the underlying profitability of a bank and possibly have an adverse impact on its credit rating. Hence, the marginal effect of the funding gap on bank ratings may depend on the overall liquidity buffers that a bank tends to hold.

We also consider measures of bank size, which is often correlated with qualitative factors that are important to bank credit analysis, such as diversification of funding sources, geographic reach, and franchise value. Moody's, for example, argues the following:

> Larger banks may often have more granular loan portfolios and broader geographic reach, reducing concentration risk. Moreover, size often allows for economies of scale, which can result in increased operating efficiency [and] may also indicate resources necessary to invest in new products and services, or to enter new markets. . . . [It] may also be an indication of greater market share, which can contribute substantially to a bank's franchise value. [Moody's Investors Service 2002b, 5]

We define bank size according to the level of total assets, where year-by-year comparisons are made by deflating total assets to constant prices. We then split banks into sample quartiles by size and define dummies for medium-small, medium-large, and large banks, using small banks as a reference category.

However, a bank that is small by global comparison may be large from a domestic perspective. That would depend, for example, on its share of lending in the local economy. Market share could then be indicative of franchise value and pricing power, correlating positively with bank ratings. In addition, banks with a higher share of lending in the economy could possibly be perceived by the market as *too important to fail*, offering them a competitive advantage in relation to funding costs, interest margins, and possibly higher ratings.[11] Higher market share could also imply higher ratings given that senior unsecured ratings by Moody's incorporate perceptions about official safety nets. As a proxy for market share, we consider the ratio of total loans by a bank to total domestic lending by the banking sector.[12] However, the sample distribution of such a ratio has a large positive skewness (10.1), meaning that some transformation is required to capture potential nonlinearities. Hence, we consider the natural logarithm of the above ratio as our variable for market share, which has a sample skewness of −0.5.

[11] O'Hara and Shaw (1990) find evidence of a positive wealth effect to large U.S. banks, resulting from the introduction of the "too big to fail" doctrine by the Comptroller of the Currency in 1984, with a corresponding negative effect on smaller banks. According to Morgan and Stiroh (1999), preferential lending terms to large U.S. banks have persisted in the 1990s even after the introduction of the Federal Deposit Insurance Corporation Act of 1991.

[12] Such a ratio tends to overstate domestic market share by internationally active banks, because its numerator includes foreign loans.

In order to control for IFRS reporting by banks, we define a relevant dummy that takes the value 1 for IFRS banks and 0 otherwise. However, the impact of IFRS reporting on modeled ratings could be ambiguous. On the one hand, IFRS could enhance comparability of financial statements across banks and help market discipline. That could eventually lead to better management, more diversified sources of funding, and, possibly, higher ratings. On the other hand, IFRS numbers could be more volatile, as discussed in Annex 1.[13] As a result, estimated ratings of IFRS banks could be lower than those of non-IFRS banks, given that any rank ordering of banks' underlying riskiness would tend to penalize banks with more volatile reported figures.

Finally, we consider a dummy to control for the dichotomy between investment- and subinvestment-grade banks (defined by the Baa3 rating threshold). Considering interaction effects between such a dummy and bank-specific variables would allow ratings of investment-grade banks to respond to shocks differently than ratings of subinvestment-grade banks. For example, it is possible that low ratings may have already factored into the possibility of more volatile financial indicators. Therefore, ratings of subinvestment-grade banks may demonstrate low sensitivity to shocks, while for investment-grade banks similar shocks could lead to a significant risk reassessment.

3.2 Country-Level Variables

Country-level variables for economic performance and banking-sector vulnerabilities could help to improve the explanatory power of the empirical model. We consider these variables both on a stand-alone basis and in the context of interaction effects (both among themselves and with bank-level variables).

As a measure of realized economic activity in a bank's country of domicile, we employ the real output gap, defined as the deviation of real GDP from a trend. As forward-looking indicators of banking-sector vulnerabilities, we consider measures of financial imbalances,

[13]The mean absolute deviation of the profitability variable is 1.2 for IFRS banks in the sample, compared with 0.8 for non-IFRS banks, which could be indicative of higher volatility in IFRS figures. This is based on 312 observations under IFRS and 1,057 under GAAP.

as suggested by Borio and Lowe (2002). In particular, we consider deviations (gaps) from a trend for the ratio of domestic credit provided by the banking sector to GDP and the stock market capitalization.[14] In order to calculate deviations of variables from their trend, we use a Hodrick-Prescott filter and annual data from 1980 to 2006 (finally revised and rebased to 100 in the year 2000).

Credit expansion above normal levels (positive gap) could be associated with periods of loose credit standards and a significant increase in credit-risk exposures by banks. As economic conditions deteriorate, credit risk could crystallize, affecting bank ratings. Similarly, a significant correction (negative gap) in the stock market could signal a change in risk perceptions about asset valuations in general, and bank assets in particular, which could possibly have an impact on bank credit ratings.[15]

In order to allow for lending boom-bust episodes to have an impact on bank credit ratings, we consider interaction effects between the real output gap and the credit gap, using various lag structures. Moreover, we consider interaction effects between the real output gap and the profitability and asset-quality variables we discussed in section 3.1. These interaction effects aim to control for cyclical variations in measures of bank profitability and asset quality, allowing their impact on bank ratings to vary with the business cycle.

As a forward-looking indicator of economic performance, we consider the slope of the yield curve. This is defined as the difference between the ten-year government bond and the three-month Treasury-bill rate. Both rates are calculated using annual averages of monthly data. According to Estrella and Hardouvelis (1991), a positive slope of the yield curve is associated with a subsequent increase in real economic activity, while a flattening of the yield

[14] Borio and Lowe also suggest that deviations of the real exchange rate from its trend may also help identify pressures building up in the capital account, as well as pressures on banks' balance sheets due to foreign exchange mismatches. However, such vulnerabilities may be more relevant to emerging-market economies with higher reliance on external capital flows and higher sensitivity to exchange rate fluctuations.

[15] For example, a fall in asset values could erode the equity buffers with which borrowers can withstand financial shocks, therefore increasing risk perceptions about secured lending. That could affect bank credit ratings if, as a result of lower collateral buffers, rating agencies thought that banks' loan portfolios had become riskier.

curve is indicative of lower future interest rates and a fall in real output. Moreover, the slope of the yield curve could correlate positively with bank profitability (especially net interest income) as a result of banks' maturity transformation role (see, e.g., Drehmann, Sorensen, and Stringa 2008).[16]

Sovereign ratings by Moody's are employed to control for sovereign credit risk in a bank's country of domicile. Sovereign ratings may act as a ceiling to senior unsecured ratings, consistent with the general-to-specific approach in producing bank ratings discussed in section 2. Moreover, senior unsecured ratings incorporate perceptions about the probability and expected scale of official safety nets. Hence, the higher the extent of official safety nets in a jurisdiction, the more we would expect bank ratings to be biased toward the sovereign ceiling. Therefore, we define dummies for sovereign ratings that correspond to the categories Aaa–Aa1, Aa2–Aa3, A1, A2, A3, Baa1–Baa3, Ba1–Ba2, and Ba3 or below, using the last category as a reference.

4. Econometric Approach

This section discusses model specification and sample selection issues. We start by describing the ordered probit approach to model bank ratings. Such an approach is particularly suitable for modeling ordinal variables, such as credit ratings, because it recognizes that the information content of one grade difference in ratings may vary along the rating scale. For example, a difference of one grade in ratings at the high end of the rating scale could imply a degree of (absolute) credit-risk differentiation that is not necessarily the same as a difference of one grade at the low end of the scale. We then discuss econometric issues relating to model and sample selection.

4.1 Ordered Probit Model

An ordered probit model of bank ratings involves the simultaneous estimation of an index variable and cut-off points for the index that

[16] According to banks' regulatory returns—for example, U.S. SEC Form 20-F—balance-sheet management and money-market revenues typically fall as a result of rising short-term interest rates and a flattening of the yield curve.

determine the transition from one rating category to another. More specifically, a number of cut-off points c_j $(j = 1, 2, \ldots, 9)$ define a time-invariant partition of index X_{it} for bank i at time t in such a way that the bank's rating R_{it+1} next period is given by

$$R_{it+1} = \begin{cases} 10 & \textit{if} \quad X_{it} \in [c_9, \infty) \\ 9 & \textit{if} \quad X_{it} \in [c_8, c_9) \\ \vdots & \qquad\qquad \vdots \\ 1 & \textit{if} \quad X_{it} \in (-\infty, c_1). \end{cases} \tag{2}$$

The unobservable index variable X_{it} is assumed to be linked to a vector V_{it} of explanatory variables through a deterministic index function $f(\cdot)$

$$X_{it} = f(V_{it}|\theta) + \varepsilon_{it}, \tag{3}$$

where θ is a vector of unknown parameters and ε_{it} is a Gaussian disturbance term with a conditional expectation of zero. For a given vector of parameters θ, cut-off points c, and explanatory variables V_{it}, the probability that bank i attains a rating R_{it+1} at time $t+1$ is given by

$$\begin{aligned} &\Pr(R_{it+1} = j|\theta, c) \\ &\quad = \begin{cases} 1 - \Phi[c_9 - f(V_{it}|\theta)] & \textit{if} \quad j = 10 \\ \Phi[c_j - f(V_{it}|\theta)] - \Phi[c_{j-1} - f(V_{it}|\theta)] & \textit{if} \quad j = 9, 8, \cdots, 2 \\ \Phi[c_1 - f(V_{it}|\theta)] & \textit{if} \quad j = 1. \end{cases} \end{aligned} \tag{4}$$

Equation (4) is estimated using maximum likelihood estimation techniques and the data panel we described in section 3 (for more details, see Greene 1997, sec. 19.8).

With the predicted probabilities from equation (4) in hand, there are various ways of predicting a bank's actual rating. Blume, Lim, and MacKinlay (1998) consider *mode* ratings (i.e., ratings with the highest probability to occur) as predictors of corporate ratings. However, if the predicted probability density is lopsided, then mode ratings may be subject to a *cliff effect*, where the most probable rating immediately follows, or precedes, a rating that has a low probability to occur. Mora (2006) considers probability-weighted (*mean*) ratings to predict sovereign ratings. Such an approach, however, could be problematic if the predicted probability density is bimodal. In that case, a mean rating would possibly be of scarce relevance because it would predict a rating that is unlikely to occur. As a modeling compromise, we use the *median* of the predicted probability density to forecast banks' actual ratings. The median rating is the one that splits the higher from the lower half of the predicted probability density of a bank's rating, rounded to the closest integer.[17]

4.2 Model Selection

In order to select the best model, we start by using a general-to-specific approach on the basis of the likelihood ratio (LR) test and the Akaike information criterion (AIC). We also examine how the model performs in and out of sample, which is our key criterion for model selection. Special emphasis is placed on the ability of the model to predict rating downgrades.

In order to select the key financial ratios described in section 3, we start from a group of candidate ratios that is our *best guess* on the basis of Moody's documentation, basic economic intuition, and the objective to avoid introducing obvious colinearity problems. If for a given CAMEL category our best guess is not statistically significant, we try alternative variables from its category and also alternative model specifications.

We also consider trade-offs between the level of sophistication of financial ratios reported by Bankscope and data availability, as well as reporting issues that could have an impact on the information content of financial ratios. For example, the tier 1 capital ratio would naturally qualify as our best guess among capital ratios.[18] Instead, we employ the ratio of shareholders' equity to total assets in order to maintain a reasonably large sample size

[17]We examined how the three approaches (i.e., mode, mean, and median ratings) perform both in and out of sample and we found that, overall, median ratings perform better than mode and mean ratings.

[18]The tier 1 capital ratio is aimed to recognize different levels of riskiness across banks and fundamental differences in bank business models.

and permit model estimation.[19] Similarly with respect to liquidity, a variable that would naturally qualify as our best guess would be the *deposit run-off ratio*. This is often defined as the ratio of liquid assets to customer deposits and short-term funds. However, under IFRS, liquid assets such as Treasury bills and other eligible bills, as well as debt securities and equity shares, are not reported separately in banks' consolidated balance sheets. Instead, they are aggregated under *trading and financial assets designated at fair value*, or *available-for-sale investments*. As a result, figures for liquid assets that are collected by Bankscope may only include a fraction of banks' actual liquid asset holdings, which could potentially lead to misleadingly low deposit run-off ratios. Therefore, we focus on bank *illiquidity*, such as our measure of funding gap described in section 3.1.

In estimating bank credit ratings, we had to consider potential endogeneity issues. Financial indicators may reflect the accessibility and price of banks' credit, as well as banks' stock market performance, which may in turn be affected by the ratings themselves. The endogeneity issue is also likely to be quite pronounced because financial indicators are observed at a low (i.e., yearly) frequency, making it difficult to establish whether these variables could have been affected by developments triggered by rating actions. We address this endogeneity issue by including all bank financial indicator variables in the model with a lag. In particular, all bank-level variables have been lagged by one year to reflect at year $t-1$ market-available information upon which our model can predict ratings in the following year t. In choosing the lags of macroeconomic variables, we tried different lag structures and we used those that we found more significant. We then confirm the lack of endogeneity problems by carrying out the Davidson and MacKinnon (1993) test for endogeneity.

White robust standard errors are used to correct for heteroskedasticity in the residuals. To adjust also standard errors for the presence of within-cluster dependence, in both the cross-section across banks and across time, we use the generalized Huber-White

[19] Defining, for example, in section 3.1 deviations from the target capital ratio in terms of the tier 1 capital ratio would significantly reduce the sample size from 1,369 observations to 1,149.

approach of Froot (1989).[20] Time-series dependence may be driven by unobserved bank effects that lead to the residuals for a given bank being correlated across time. Unobserved bank effects may result from qualitative factors as well as from different interpretation of accounting policies, which may affect the information content of financial ratios across jurisdictions. Cross-sectional dependence implies that the residuals for a given year are correlated across banks. That could result from broad changes in accounting policies, such as the IFRS transition, and the implementation of new prudential standards. Industry-wide trends may also give rise to cross-sectional correlation as a result of developments in both the asset side (e.g., credit expansion) and the liability side (e.g., funding gap) of banks' balance sheets.

5. Empirical Results

The estimated coefficients of the ordered probit model are reported in table 4. Overall, bank-specific variables are statistically significant and consistent with economic intuition and public statements by Moody's. An important result is that we find significant evidence of procyclicality in bank ratings that manifests itself through lagged interaction effects between the real output and credit gap. But we also find evidence that bank ratings internalize cyclical variations in asset quality by penalizing low asset quality more aggressively in good times than in periods of economic slowdown. However, no significant evidence is found that bank ratings internalize cyclical variations in profitability, which could be a potential source of procyclicality in ratings. Bank ratings also appear to respond positively to a steepening of the yield curve and also tend to increase with sovereign ratings.

Next we discuss our empirical results in more detail. It should be emphasized that financial ratios, as well as gap and interest rate variables, are expressed in percentage terms, which may result in some coefficients appearing small in absolute terms, although economically significant. Therefore, the far-right column of table 4 reports the coefficient for each variable multiplied with the corresponding

[20] For a description of how standard errors are adjusted for within-cluster correlation in Stata, see Rogers (1993).

Table 4. Ordered Probit Model for Bank Senior Unsecured Ratings by Moody's, for the Period 1999–2006

		Investment-Grade Bank		Interaction Effects if Subinv.-Grade Bank		
DEPENDENT VARIABLE: *Bank Senior Unsecured Rating by Moody's*	Number of Lags (in Years)	Coefficient	Robust Std. Error	Coefficient	Robust Std. Error	Coefficient × (Mean Abs. Deviation)
INDEPENDENT VARIABLES:						
Bank-Level Financial Indicators						
Capital: Deviation of (Equity/Total Assets) Ratio from Target *c1*	1	0.007***	0.002	−0.004**	0.002	0.109
c2	1	0.017*	0.009	0.015	0.024	0.100
c3	1	−0.002	0.003	0.002	0.003	−0.001
Asset Quality: Loan Loss Reserves/Net Interest Income	1	−0.005**	0.002	0.000	0.003	−0.084
*Asset Quality*Real GDP Gap*	1	−0.056***	0.021	—	—	−0.041
Cost Efficiency: Operating Costs/Total Assets	1	−0.084***	0.032	−0.163*	0.091	−0.205
Recurring Earning Power: Pre-Tax, Pre-Prov. Profits/Total Assets	1	0.194***	0.071	−0.178**	0.079	0.043
*Recurring Earning Power*Domestic Business Cycle*	1,1	−0.027	0.189	—	—	−0.002
Funding Gap: (Loans – Deposits)/Loans	1	−0.002***	0.000	0.007***	0.002	−0.142
Market Share: Log (Loans/Domestic Credit Extension by Banks)	1	0.213***	0.047	—	—	0.393
Bank-Level Dummies						
Subinvestment-Grade Bank: 1 if bank rating below Baa2–Baa3	1	—	—	−2.236***	0.409	—
IFRS Bank: 1 if consolidated accounts in IFRS	1	−0.631***	0.121	—	—	—
Bank Size:						
Large Bank: 1 if in 4th quartile by total assets ($US)	1	1.815***	0.195			—
Medium-Large Bank: 1 if in 3rd quartile by total assets ($US)	1	1.074***	0.162			—
Medium-Small Bank: 1 if in 2nd quartile by total assets ($US)	1	0.637***	0.129			—

(continued)

Table 4. (Continued)

Independent Variables:	Number of Lags (in Years)	Investment-Grade Bank Coefficient	Investment-Grade Bank Robust Std. Error	Interaction Effects if Subinv.-Grade Bank Coefficient	Interaction Effects if Subinv.-Grade Bank Robust Std. Error	Coefficient × (Mean Abs. Deviation)
Macroeconomic Variables						
Yield-Curve Slope: (10-year gov. bond rate) – (3-month T-bill rate)	1	0.051*	0.028			0.043
Domestic Business Cycle: Real GDP gap	1	0.020*	0.011			0.052
Domestic Bank Credit Cycle: Bank credit gap	2	−0.001	0.002			−0.013
*Domestic Bank Credit Cycle*Domestic Business Cycle*	2,1	0.169**	0.087			0.037
Stock Market Performance: Market capitalization gap	1	−0.001	0.001			−0.018
Sovereign Rating Dummies						
Aaa–Aa1	1	4.560***	0.443			
Aa2–Aa3	1	4.484***	0.457			
A1	1	2.328***	0.462			
A2	1	2.194***	0.464			
A3	1	1.901***	0.412			
Baa1–Baa3	1	1.823***	0.415			
Ba1–Ba2	1	1.744***	0.375			
Region Dummies						
Denmark, Finland, Norway, Sweden	—	0.587***	0.211			
Germany	—	0.429**	0.187			
United Kingdom	—	0.394***	0.130			
Australia	—	−0.391**	0.189			
Italy	—	−0.525***	0.190			
Japan	—	−1.198***	0.161			

(continued)

Table 4. (Continued)

Independent Variables:	Number of Lags (in Years)	Investment-Grade Bank Coefficient	Investment-Grade Bank Robust Std. Error	Interaction Effects if Subinv.-Grade Bank Coefficient	Interaction Effects if Subinv.-Grade Bank Robust Std. Error	Coefficient × (Mean Abs. Deviation)
Lower Boundary for Rating Categories						
Aaa–Aa1	—	6.929	0.597			
Aa2	—	6.402	0.586			
Aa3	—	5.489	0.588			
A1	—	4.725	0.585			
A2	—	3.746	0.579			
A3	—	2.752	0.583			
Baa1	—	1.188	0.582			
Baa2–Baa3	—	−0.851	0.608			
Ba1–Ba3	—	−3.061	0.548			
B1 and Below	—	$-\infty$	—			

Note: The model is estimated using a data panel of 1,369 observations, for the period 1999–2006. The data panel includes published accounts data of 293 banks from thirty-three countries (grouped in fourteen regions) and macroeconomic information. The original sample consisted of 2,022 observations. Of those, 251 observations were dropped due to lags in the explanatory variables; 350 observations were also omitted due to lack of contemporaneous observations for all explanatory variables; and 52 observations were omitted because of rating withdrawals.

*Significant at 10 percent; **significant at 5 percent; ***significant at 1 percent. The first column reports the lag structure of explanatory variables in the model. For interaction effects we report two lags, one for each interacting variable. The second column reports the estimated coefficients of explanatory variables in the model. The third column reports robust standard errors. The fourth column reports the estimated coefficients of interaction effects between explanatory variables and a dummy that takes the value 1 if the bank had a previous-year rating of subinvestment grade (below Baa2–Baa3) and 0 otherwise. Robust standard errors for these interaction effects are reported in the fifth column. The last column reports the product of the estimated coefficient with the corresponding mean absolute deviation of each explanatory variable (excluding dummies).

mean absolute deviation, which could help to assess the economic significance of estimated coefficients.

5.1 Bank-Level Variables

Starting with bank-level variables, the coefficients of the first and second transformation of the capital variable (*c1* and *c2*) are statistically significant and with the expected sign. Negative deviations from the capital target (the second transformation) appear to have a larger absolute impact on ratings than positive deviations of equal magnitude (the first transformation). Also, the coefficient for positive deviations from the capital target appears more economically significant for investment-grade than for subinvestment-grade ratings. This is not surprising given that earnings of low-rated banks may be weaker to absorb adverse shocks, meaning that higher capital buffers would be needed to protect investors.[21] Hence, from a Moody's perspective, banks of lower credit quality and weaker profitability would possibly need to hold higher capital buffers to support their ratings, which could lessen any positive impact on ratings from holding capital above a theoretical target.

The coefficient for the ratio of loan loss reserves to net interest income is statistically significant and with a negative sign. The less net interest margins compensate for risks in loan portfolios, the lower is asset quality and the more negative the impact on ratings. But what is more interesting is that bank ratings tend to penalize low asset quality more aggressively when the economy is booming (positive real output gap) than when it is slowing down (negative real output gap). This is reflected in the statistically significant and negative coefficient for interaction effects between the real output gap and asset quality. Given that measures of asset quality tend to

[21]For subinvestment-grade banks, the sample distribution of the profitability variable has a mean absolute deviation of 1.6, compared with 0.7 for investment-grade banks. This is in line with intuition that earnings of low-rated banks are more volatile and, hence, less reliable to withstand adverse shocks in the long run. In addition, the sample distribution of the asset-quality variable has a mean absolute deviation of 34 for subinvestment-grade banks, compared with 14 for investment-grade banks, which implies that low-rated banks may be subject to larger shocks.

be procyclical (see, e.g., Albertazzi and Gambacorta 2006; Lown, Morgan, and Rohatgi 2000), the above result indicates that bank ratings tend to respond to asset-quality changes in a countercyclical fashion.

Less cost-efficient banks tend to attract lower ratings. This is reflected in the negative and statistically significant coefficient for the ratio of operating costs to total assets. Cost efficiency appears to be a particularly important driver of subinvestment-grade ratings. This is reflected in the statistically significant and negative coefficient for the interaction effect between the cost-efficiency ratio and the subinvestment-grade dummy. A similar result is obtained if, as a measure of cost efficiency, we use the ratio of operating costs to income.[22]

Profitability has a significant impact on bank ratings, as indicated by the statistically significant and positive coefficient for recurring earning power. Profitability also appears to be a more important driver of investment- than subinvestment-grade ratings, as reflected in the negative coefficient for interaction effects between recurring earning power and the subinvestment-grade dummy. This result supports the hypothesis that higher uncertainty about earnings may have already been incorporated into lower ratings which, as a result, may show lower sensitivity to shocks in earnings than investment-grade ratings. However, we find no evidence that ratings internalize cyclical variations in bank profitability. In particular, the coefficient for interaction effects between recurring earning power and real output gap is not statistically significant, although it appears with a negative sign. A negative coefficient would be indicative of a *discount* in the rating process for profits during good times, or a *premium* during periods of economic slowdown.

Regarding banks' liquidity position, investment-grade banks may get easy access to external sources of funding, which could reduce their marginal propensity to hoard (low-yielding) liquid assets. Then a higher funding gap would imply a higher ratio of (illiquid) customer

---

[22] In that case, the estimated coefficient for cost efficiency of investment-grade banks is not statistically different than 0, while for subinvestment-grade banks the coefficient (-0.011) is statistically significant at 5 percent. This further supports the hypothesis that cost efficiency is an important determinant of subinvestment-grade ratings.

loans to short-term liabilities and, hence, a lower buffer of liquid assets against unanticipated foreclosures of credit lines. That could lead to lower ratings, as indicated by the negative and statistically significant coefficient for funding gap in table 4. But subinvestment-grade banks may already hoard too much liquidity to finance future investment if access to wholesale funding is relatively costly. High levels of liquidity hoarding could then imply inefficient employment of financial resources, lower future profitability, and possibly lower ratings, as indicated by the positive coefficient for funding gap.

Larger banks also tend to attract higher ratings, given that all coefficients for size dummies are statistically significant and monotonic. As discussed in section 3.1, this is in line with the intuition that bank size could correlate with cross-border diversification of assets and funding sources, internal economies of scale, and, possibly, management quality. Moreover, the coefficient for domestic market share of lending is significant and with a positive sign. This is in line with the intuition that market share could be indicative of domestic franchise value, pricing power, and, possibly, systemic importance—all of which could lead to higher ratings.

Finally, we find evidence of a negative bias in the ratings of IFRS reporting banks, given the statistically significant and negative coefficient for the IFRS dummy. As discussed in section 3.1, such a negative bias could be due to higher volatility in IFRS numbers. Also, the benefits of IFRS reporting (e.g., better-quality management due to market discipline) may take some time to reflect in actual ratings and may not be fully captured in the early years of IFRS implementation that we consider in the sample.

5.2 Country-Level Variables

Turning to macroeconomic variables, the lagged interaction effect between the credit gap and the real output gap is statistically significant and with a positive sign, which is indicative of procyclicality in bank ratings. When a credit boom (positive credit gap, in two-year lag) is followed by a subsequent macroeconomic slowdown (negative real output gap, in one-year lag), current ratings tend to be lower. However, when credit expansion above normal levels is followed by a year of economic boom, bank credit ratings tend to increase. A similar result is obtained if, instead of the real output gap, we consider a dummy for economic slowdowns (taking the value 1, if real output gap is negative, and 0 otherwise) and its interaction with the credit gap.[23] In addition, the coefficient for real output gap is statistically significant and positive, adding further support to the hypothesis of procyclicality in bank ratings.

A steepening of the yield curve also appears to have a positive impact on bank ratings, as indicated by the positive and significant coefficient for the yield-curve slope. This result is consistent with empirical evidence that identifies the slope of the yield curve as a predictor of turns in the business cycle (Estrella and Hardouvelis 1991). It is also consistent with banks' asset-liability repricing mismatch (implied by their maturity transformation function), which makes banks susceptible to a flattening of the yield curve (Drehmann, Sorensen, and Stringa 2008).

Finally, all dummies for sovereign ratings are statistically significant and monotonic. Ceteris paribus, the higher the sovereign rating of a bank's country of domicile, the higher the bank's rating. Country effects have possibly been absorbed by macroeconomic variables, country ratings, and the constructed variable to capture possible deviations from the capital target, which also considers country effects. Hence, most of the regional dummies in the estimated model are not significant, with table 4 presenting the coefficients of the statistically significant ones. These are in line with our prior about the level of banking-system development, the existence of state-owned institutions in the sample, and government guarantees. Scandinavian banks appear to benefit the most from the country effect, followed by German and UK banks. Japan has the lowest coefficient, probably reflecting the problems experienced in the Japanese banking system over the past decade, which nevertheless have not impacted materially on Japan's sovereign rating.[24]

[23]In line with the previous specification, the interaction effect in that case has a negative and statistically significant coefficient. Estimation results relating to alternative specifications of the model are available from the authors upon request.

[24]There is no apparent connection between the ratings of Japanese banks and the sovereign rating of Japan. While the Japanese banking system was in crisis, Japan was rated Aaa by Moody's until November 16, 1998, when it was downgraded to Aa1. It was upgraded again to Aaa on October 20, 2002.

5.3 Goodness-of-Fit and Robustness Checks

In order to assess the performance of the estimated model that is presented in table 4, we compare model predictions with actual ratings by Moody's. To illustrate such a comparison, we use bar charts showing the proportion of actual ratings that are correctly predicted by the model, as opposed to ratings that are either over- or underpredicted. For expositional convenience, comparisons between actual and predicted ratings are presented in terms of high (Aaa–Aa3), medium (A1–A3), and low (Baa1–Caa3) rating categories. We also show the total results across rating categories.

In-sample prediction results are presented in figure 1. Overall, the model predicts correctly 45 percent of actual ratings in the sample. In terms of model performance across rating categories, the model predicts correctly 41 percent of high ratings, 44 percent of medium ratings, and 52 percent of low ratings.[25] The number of over- and underpredictions by the model are almost equally balanced at 28 percent and 27 percent, respectively. Table 5 offers a more detailed analysis of the goodness of fit of the estimated model. For example, the table shows that from 212 actual Aa3 ratings, the model predicts correctly 122, while it overpredicts 33 (by assigning a rating Aaa–Aa1 to 6 and Aa2 to 27) and underpredicts 57 (by assigning a rating A1 to 41, A2 to 13, and A3 to 3).

We observe that the estimated model tends to underpredict high ratings and overpredict low ratings. Figure 1 shows that for ratings above Aa3, the incidence of underprediction is 50 percent, compared with 9 percent of overprediction. Yet, for ratings between Baa1 and Caa3, the incidence of underprediction falls to 10 percent, while overprediction errors increase to 38 percent. For medium ratings, model errors are more balanced, representing 21 percent underprediction and 35 percent overprediction. To some extent, such a bias toward underrating high ratings and overrating low ratings is imposed by the model structure. In other words, the only way to err in predicting a rating at the top end of the rating spectrum is to underpredict it, and vice versa for the bottom end.

[25] From the 1,369 data points considered, 411 correspond to banks rated Aaa–Aa3, 622 to banks rated A1–A3, and 336 to banks rated Baa1–Caa3.

Figure 1. In-Sample Estimates (Median Ratings)

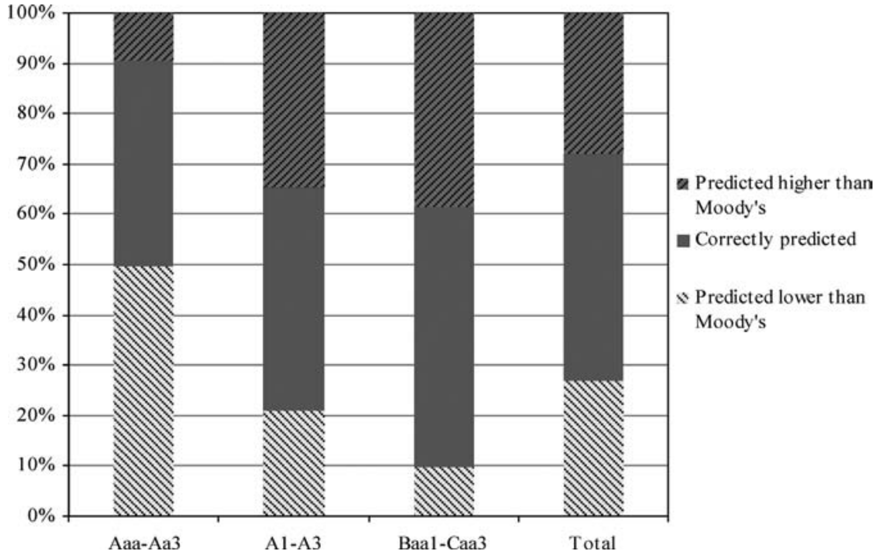


Table 5. Ordered Probit Model Predictions of Actual Ratings, for the Period 1999–2006

	Predicted Ratings										
Actual Ratings	Aaa–Aa1	Aa2	Aa3	A1	A2	A3	Baa1	Baa2–Baa3	Ba1–Ba3	B1 or Below	Total
Aaa–Aa1	11	16	49	24	2	2	0	0	0	0	104
Aa2	6	35	34	15	4	1	0	0	0	0	95
Aa3	6	27	122	41	13	3	0	0	0	0	212
A1	0	5	36	89	55	14	2	0	0	0	201
A2	0	4	23	62	108	41	5	0	0	0	243
A3	0	0	4	8	74	79	10	3	0	0	178
Baa1	0	0	0	4	14	77	10	10	1	0	116
Baa2–Baa3	0	0	0	0	1	1	4	76	17	0	99
Ba1–Ba3	0	0	0	0	0	0	0	22	58	5	85
B1 or Below	0	0	0	0	0	0	0	1	5	30	36
Total	23	87	268	243	271	218	31	112	81	35	1, 369

Note: The matrix shows the number of actual versus predicted ratings using the estimated ordered probit model presented in table 4. Diagonal elements represent the number of correct model predictions per actual rating category on the far-left column. Elements above the diagonal represent number of underpredictions per actual rating category, while elements below the diagonal represent number of overpredictions.

The goodness of fit of the model at the top end of the rating spectrum may also be affected by omitted variables. We recognize that with the exception of bank size, market share, and country effects, the model does not consider variables for geographic and sectoral diversification, risk-management expertise, quality of staff, and integrity. In addition, it does not explicitly control for banks in the sample that are state sponsored (i.e., state-owned banks, or banks that are covered by government guarantees). Such banks would possibly receive the sovereign (ceiling) rating, regardless of their underlying financial indicators. Yet some of the impact of state sponsorship on bank ratings may be already captured by explanatory variables, such as country effects, market share, and dummies for sovereign ratings, as discussed in section 3.1.[26]

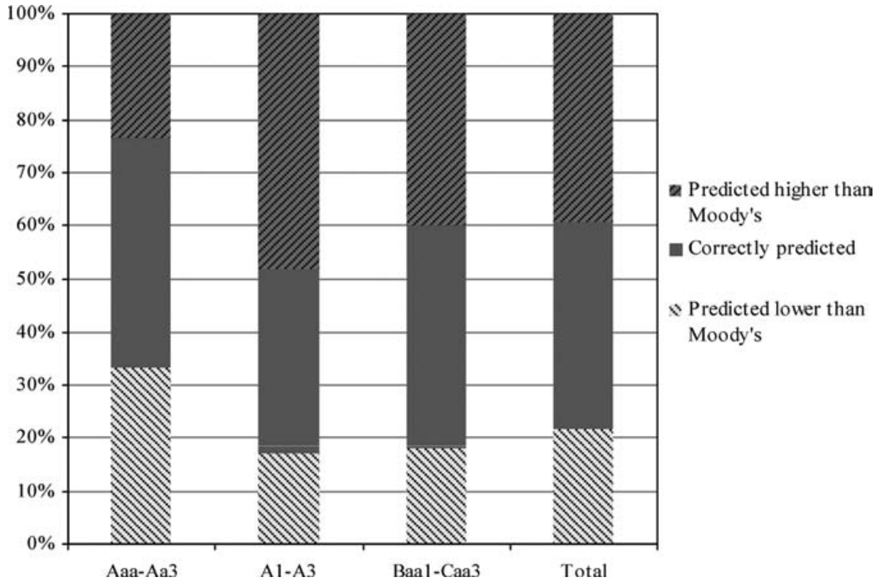
Another factor that could lead to underprediction errors for high ratings is the way that predicted ratings are defined. As discussed in section 4.1, we consider the median of the estimated probability distribution of ratings as the predicted rating. But for the median to correctly predict the rating category Aaa–Aa1 would require the estimated probability for that category to be at least 50 percent. By considering the most probable (mode) rating, we are able to improve the model performance in predicting Aaa–Aa1 ratings. In particular, compared with eleven correct predictions of Aaa–Aa1 ratings under a median-rating approach (see table 5), mode ratings predict correctly twenty-five ratings. However, that comes at a cost of higher prediction errors at lower rating categories, compared with median ratings.

The definition of the rating categories (see section 3) may also affect the goodness of fit of the model. We tried alternative designs of the rating categories (maintaining ten rating buckets), and the goodness of fit of the model remains broadly unchanged. But compared with other studies of credit ratings, the number of rating categories (ten) that we consider is relatively large.[27] As a result, the predicted

[26]For example, approximately one-third of the banks rated Aaa–Aa1 in the sample are German banks, especially Landesbanks, whose debt issuance until July 2005 was covered by explicit state guarantees. As already discussed in section 5, ratings of German banks are among those that benefit most from the country effect.

[27]Amato and Furfine (2004) consider eight rating categories and Blume, Lim, and MacKinlay (1998) only four, with success rates in predicting ratings 53 percent and 57 percent, respectively.

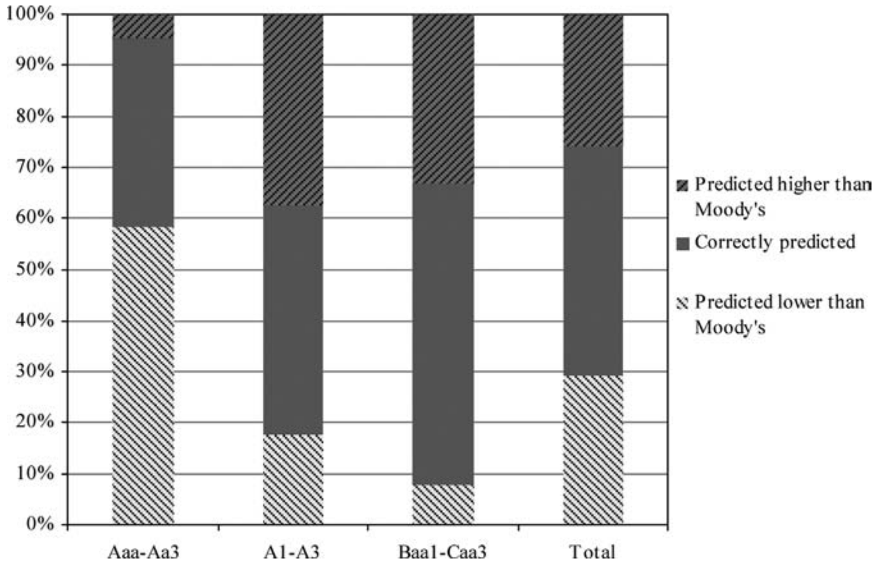
Figure 2. Out-of-Sample Estimates for 2005 (Median Ratings)



probabilities are spread across a relatively wide range of ratings categories, which makes it harder for median ratings to predict actual ratings. Therefore, we also considered coarser rating buckets, and the goodness of fit of the predicted model substantially improved. However, attaining a better fit of the model using coarser rating buckets would come at a cost of worse predictions relative to the true Moody's scale. That would clearly limit the practical use of the empirical model and, hence, we prefer more finely defined rating buckets. In any case, by grouping in table 5 the predicted and actual ratings into four rating categories (Aaa–Aa3, A1–A3, Baa1–Baa3, and B1 or below) similar to Blume, Lim, and MacKinlay (1998), 1,064 out of the 1,369 data points are "correctly predicted," which implies a quasi-success rate of approximately 78 percent.

We also consider how the model performs out of sample for the years 2005 and 2006. Out-of-sample predictions for 2005 are based on six years of data from 1999 through 2004 and are presented in figure 2. Predictions for 2006 are based on seven years of data from

Figure 3. Out-of-Sample Estimates for 2006 (Median Ratings)

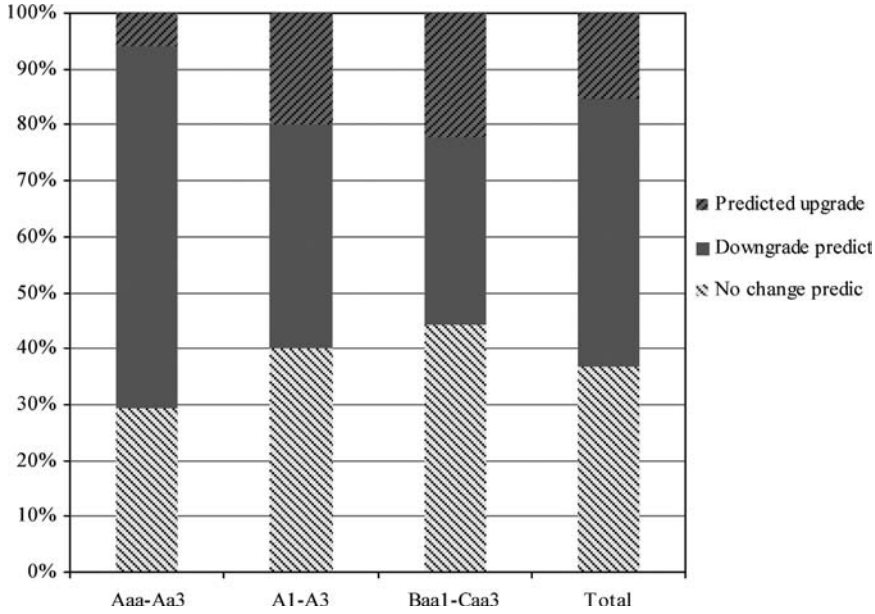


1999 through 2005 and are shown in figure 3. The estimated model predicts correctly 39 percent of ratings in 2005 and 45 percent of ratings in 2006.[28] Finally, we examine how well the model performs in predicting ratings downgrades. As figure 4 shows, the model predicts correctly twenty-two out of forty-six downgrades (48 percent), while in seven cases it predicts an upgrade, and in seventeen cases it predicts no change in ratings. The highest proportion of correctly predicted downgrades is achieved for the Aaa–Aa1 rating category (65 percent).[29]

[28]For the year 2005, we consider 188 observations, with 51 for high ratings (Aaa–Aa3), 87 for medium ratings (A1–A3), and 50 for low ratings (Baa1–Caa3). For 2006 we consider 195 observations, with 65 for high ratings, 91 for medium ratings, and 39 for low ratings.

[29]For the Aaa–Aa1 rating category, the model predicts eleven out of seventeen downgrades (65 percent); for A1–A3 it predicts eight out of twenty downgrades (40 percent); and for Baa1–Caa3 the model predicts three downgrades out of nine (33 percent).

Figure 4. Predicting Downgrades (Median Ratings)



6. Conclusions

A number of studies suggest that the financial system is intrinsically procyclical (e.g., Bernanke, Gertler, and Gilchrist 1999). Hence, credit risk that is built up on banks' balance sheets during good times may crystallize as credit and economic conditions deteriorate. Downgrades in bank ratings could then feed into negative market sentiment about the banking sector, precipitating a deleveraging process by banks attempting to improve their financial indicators. That could feed into a cycle of further tightening of credit conditions, financial distress by borrowers, and deterioration in banks' financial indicators and ratings. Therefore, a closer examination of the behavior of bank ratings and, in particular, of possible channels through which procyclicality in ratings could manifest itself, would be of interest both to market participants and to policymakers. To our knowledge, this is the first paper that discusses procyclicality of bank credit ratings focusing on (lagged) interaction effects between the credit and business cycle. It also examines the extent to which bank ratings internalize cyclical variations in measures of asset quality and profitability.

Controlling for bank financial characteristics, we find evidence that bank senior unsecured ratings correlate positively with the slope of the yield curve, sovereign ratings, market share of lending, and bank size. Moreover, we find significant evidence of procyclicality in bank ratings owing to lead-lag interaction effects between the real output gap and the credit gap. This is consistent with evidence from Moody's that changes in corporate bond *ratings are strongly correlated with cyclical indicators such as economic activity, default rates, and credit spreads* and that *average rating levels generally move in tandem with the cycle* (see Cantor, Mahoney, and Mann 2003). Bank ratings also appear to internalize cyclical variations in asset quality by penalizing low asset quality more aggressively in good times than in periods of economic slowdown. However, no significant evidence is found that bank ratings distinguish between profitability at different stages of the business cycle.

Bank ratings could correlate with the credit and economic cycle as a result of difficulties faced by market participants (including rating agencies) in assessing how systemwide risks evolve over time, or distinguishing between cyclical variations and structural changes. Such signal extraction problems could be exacerbated by complex feedback effects between the financial and the real sector, product innovation, and evolution of business models by banks. Although an analysis of these issues is beyond the scope of this paper, our results indicate that procyclicality in bank ratings could possibly be mitigated by adjusting the degree of pass-through of earnings performance into ratings, conditional on the stage of the economic and credit cycle.

Appendix

IFRS Reporting and Modeling Implications

IFRS are aimed to offer a more realistic picture of profits and losses due to full disclosure of income and costs that arise, for example, from insurance business and the fair-value treatment of certain assets (see Bank of England 2005, 42). IFRS could also facilitate cross-border comparisons of financial statements and, through stricter disclosure standards, could increase market discipline.[30] That could increase management efficiency and enhance the diversification of funding sources, which could potentially lead to higher bank ratings.

However, IFRS could also lead to higher volatility of reported figures, both across time and in the cross-section across banks. Reported figures, for example, could appear more volatile under IFRS as a result of the fair-value option in accounting for financial instruments and off-balance-sheet items (IAS 39). Under IFRS, such a fair-value option is combined with *neutrality*, which could lead to less smoothing of financial results over time. This represents a departure from many local GAAP standards, where income and expense are calculated on an accrual basis, financial instruments are accounted at historical cost (unless qualified for inclusion in the trading book), and there is a conservative bias toward *prudence* embedded in the accounts. Similarly, IFRS banks are prevented from provisioning against bad loans on a forward-looking basis, which could induce further procyclicality in their financial results (IAS 37). Last but not least, IFRS is a *principles-based* framework that, according to market commentators, could offer more leeway for interpretation, compared with well-developed rules-based systems, such as the US GAAP.[31] That could lead to a wider set of results under IFRS reporting, higher implementation uncertainties, and, possibly, higher litigation risk due to lawsuits by investors.[32]

More than 100 countries—including all EU countries, Australia, Canada, China, Japan, and Russia—are now using or adopting IFRS. Under EU regulation, all listed companies, including banks, are required to produce their consolidated financial statements according to IFRS, beginning January 2005. The majority of EU banks restated their 2004 financial results under IFRS to permit consistent computation and comparison of growth rates. Banks may

[30]IFRS 7, for example, requires companies to make adequate disclosure about judgments and uncertainties in valuing financial instruments.

[31]As an indication of the potential scope for interpretation under IFRS, the IFRS principles-based framework is covered in some 2,500 pages, while the U.S. GAAP rules-based system is described in more than 25,000 pages.

[32]See, e.g., "A Single Standard for the World?" *Financial Times,* March 25, 2008.

also opt for IFRS reporting, alongside their national GAAP numbers, regardless of regulatory requirements to do so. Despite the fact that IFRS and U.S. GAAP have moved closer together since 2002,[33] it is only since 2006 that FASB and the IASB have agreed to a formal plan of convergence between the two sets of standards.[34] Given that the sample covers the period 1999–2006 (i.e., before the formal inauguration of convergence between IFRS and U.S. GAAP), the two sets of rules are considered distinct for the purposes of our analysis.

References

Albertazzi, U., and L. Gambacorta. 2006. "Bank Profitability and the Business Cycle." Bank of Italy Economic Research Paper No. 601.

Amato, J. D., and C. H. Furfine. 2004. "Are Credit Ratings Procyclical?" *Journal of Banking and Finance* 28 (11): 2641–77.

Bank of England. 2005. "The Financial Stability Conjuncture and Outlook." *Financial Stability Review* December (19): 7–46.

Bernanke, B. S., M. Gertler, and S. Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework." In *Handbook of Macroeconomics*, ed. J. B. Taylor and M. Woodford, vol. 1, 1341–93. Amsterdam: North Holland.

Blume, M. E., F. Lim, and A. G. MacKinlay. 1998. "The Declining Credit Quality of U.S. Corporate Debt: Myth or Reality?" *Journal of Finance* 53 (4): 1389–1413.

Borio, C., and P. Lowe. 2002. "Assessing the Risk of Banking Crises." *Bank for International Settlements Quarterly Review* (December): 43–54.

Cantor, R., C. Mahoney, and C. Mann. 2003. "Are Corporate Bond Ratings Procyclical?" Moody's Special Comment (October).

[33]For example, the Norwalk agreement in October 2002 between the U.S. Financial Standards Board (FASB) and the International Financial Standards Board (IASB) resulted in the first common standard (IFRS 5), and in March 2003 FASB recognized the need for a *principles-based* approach to standard setting, similar to IFRS.

[34]In February 2006, the two Boards signed a memorandum of understanding that laid down a roadmap of convergence between U.S. GAAP and IFRS in the period from 2006 through 2008.

Cantor, R., and C. Mann. 2007. "Analyzing the Trade-off between Rating Accuracy and Stability." *Journal of Fixed Income* (Spring): 60–68.

Davidson, R., and J. MacKinnon 1993. *Estimation and Inference in Econometrics.* Oxford: Oxford University Press.

Dell' Ariccia, G., D. Igan, and L. Laeven. 2008. "Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market." IMF Working Paper No. 106.

Diamond, D., and P. Dybvig. 1983. "Bank Runs, Deposit Insurance and Liquidity." *Journal of Political Economy* 91 (3): 401–19.

Drehmann, M., S. Sorensen, and M. Stringa. 2008. "The Integrated Impact of Credit and Interest Rate Risk on Banks: An Economic Value and a Capital Adequacy Perspective." Bank of England Working Paper No. 339.

Estrella, A., and G. Hardouvelis. 1991. "The Term Structure as a Predictor of Real Economic Activity." *Journal of Finance* 46 (2): 555–76.

Fight, A. 2001. *The Ratings Game.* Wiley & Sons.

Flannery, M. J., S. H. Kwan, and M. Nimalendran. 2004. "Market Evidence on the Opaqueness of Banking Firms' Assets." *Journal of Financial Economics* 71 (3): 419–90.

Froot, K. A. 1989. "Consistent Covariance Matrix Estimation with Cross-Sectional Dependence and Heteroskedasticity in Financial Data." *Journal of Financial and Quantitative Analysis* 24 (3): 333–55.

Greene, W. H. 1997. *Econometric Analysis.* 3rd ed. Englewood Cliffs, NJ: Prentice-Hall.

Jiménez, G., and J. Saurina. 2006. "Credit Cycles, Credit Risk and Prudential Regulation." *International Journal of Central Banking* 2 (2): 65–98.

Lown, C., D. Morgan, and S. Rohatgi. 2000. "Listening to Loan Officers: The Impact of Commercial Credit Standards on Lending and Output." *Economic Policy Review* (Federal Reserve Bank of New York) 6 (2): 1–16.

Moody's Investor Service. 1999. "Bank Credit Risk in Emerging Markets." New York.

———. 2002a. "An Update to the Key Ratios Moody's Uses in Its Analysis of Finance Companies." New York.

———. 2002b. "Ratios and US Bank Ratings." New York.

———. 2006a. "Bank Financial Strength Ratings: Revised Methodology." New York.

———. 2006b. "Banking System Outlook: United Kingdom." New York.

Mora, N. 2006. "Sovereign Credit Ratings: Guilty beyond Reasonable Doubt?" *Journal of Banking and Finance* 30 (7): 2041–62.

Morgan, D. P. 2002. "Rating Banks: Risk and Uncertainty in an Opaque Industry." *American Economic Review* 92 (4): 874–88.

Morgan, D. P., and K. J. Stiroh. 1999. "Bond Market Discipline of Banks: Is the Market Tough Enough?" Federal Reserve Bank of New York Staff Report No. 95.

O'Hara, M., and W. Shaw. 1990. "Deposit Insurance and Wealth Effects: The Value of Being 'Too Big to Fail'." *Journal of Finance* 45 (5): 1587–1600.

Rogers, W. H. 1993. "Regression Standard Errors in Clustered Samples." *Stata Technical Bulletin* 13: 19–23. Reprinted in *Stata Technical Bulletin Reprints* 3: 88–94.

Shin, H. S., and T. Adrian. 2007. "Liquidity and Leverage." Paper presented at the Sixth BIS Annual Conference on Financial System and Macroeconomic Resilience, Brunnen, Switzerland, June 18–19.

Explaining Monetary Policy in Press Conferences*

Michael Ehrmann and Marcel Fratzscher
European Central Bank

The question of how best to communicate monetary policy decisions remains a highly topical issue among central banks. Focusing on the experience of the European Central Bank, this paper studies how explanations of monetary policy decisions at press conferences are perceived by financial markets. The empirical findings show that ECB press conferences provide substantial additional information to financial markets beyond that contained in the monetary policy decisions, and that the information content is closely linked to the characteristics of the decisions. Press conferences have on average had larger effects on financial markets than the corresponding policy decisions, with lower effects on volatility. Moreover, the Q&A part of the press conference fulfills a clarification role, in particular during periods of large macroeconomic uncertainty.

JEL Codes: E52, E58, G14.

1. Introduction

The way central banks communicate with the public has seen dramatic changes in recent decades. Further modifications to current practices are in the making, with a number of central banks currently debating whether and how to modify communication practices. An

*We would like to thank seminar participants at the ECB, as well as Klaus Adam, Helge Berger, Niels Bünemann, Luca Dedola, Carlo Rosa, Frank Smets, Rolf Strauch, Philip Vermeulen, an anonymous referee, and especially Hyun Shin for comments on earlier versions of the paper, Andreas Strauch for continuous and insightful discussions of the financial market's perception of the ECB's communication, and Alexander Bick for excellent research assistance. The views expressed in this paper are those of the authors and do not necessarily reflect those of the European Central Bank. Author contact: Michael.Ehrmann@ecb.int and Marcel.Fratzscher@ecb.int; European Central Bank, Kaiserstrasse 29, D-60311 Frankfurt/Main, Germany.

interesting case in point is the decision in 2007 by the Swedish Riksbank to hold a press conference after each policy meeting to "provide more detailed and more regular information." At the time, the press conference had been given substantially more weight in the Riksbank's communication strategy, as it was initially intended (yet changed in 2008) not to give any further communication on monetary policy intentions through speeches in the intermeeting periods.[1]

Overall, there is a clear tendency to provide more information, and to do so in a much more timely fashion. Not only do central banks communicate more about their policy objectives (most prominently in the case of inflation targets) and their strategies, but the way central banks communicate their monetary policy decisions has also evolved considerably. In general, enhanced communication and transparency is widely argued by both policymakers and academics to have improved the effectiveness of monetary policy considerably.[2]

Whereas it is nowadays common practice to announce policy decisions immediately by means of a press release, central banks have adopted various approaches as to how policy decisions are *explained* to the public. One relatively recent approach has been the introduction of press conferences, where monetary policy decisions are explained in detail and journalists are given the chance to ask questions to the central bank officials.[3] Regular press conferences to explain monetary policy decisions are currently held by the central banks of the Czech Republic, Japan, New Zealand, Norway,

[1]On May 11, 2007, First Deputy Governor Irma Rosenberg declared the following: "Firstly, press conferences will in future be held after each monetary policy meeting, regardless of what decision has been taken. . . . By [. . .] holding press conferences after each monetary policy meeting the Riksbank will provide more detailed and more regular information on the considerations taken by the Executive Board." Moreover, she stated, "The Executive Board has come to the conclusion that there is not normally any reason to indicate how the repo rate will be set in speeches and press releases issued prior to the monetary policy meetings. Our assessment is that it is enough to signal our intentions clearly in connection with the seven monetary policy meetings held every year." (Rosenberg 2007)

[2]This point is stressed by a number of important studies—though this list is by no means exhaustive—including Blinder (1998), Bernanke (2004), Goodhart (2005), Issing (2005), Woodford (2005), and Reinhart and Sack (2006).

[3]With the notable exceptions of the Swedish and Swiss central banks, the introduction of regular press conferences dates back only to the turn of the millennium (Issing 2005).

Poland, Sweden, and Switzerland, as well as by the European Central Bank (ECB). An alternative approach has been to provide only a short statement on the decision on the meeting day, followed by the release of minutes, usually a few weeks later. This approach is currently employed in particular by the Bank of England and the Federal Reserve.

With a view to the ongoing reassessment of communication strategies of central banks, and having gained some experience with press conferences as a communication instrument, a first evaluation of their usefulness is now in order. This paper analyzes the case of the ECB's press conferences, which have been part of the ECB's communication tools right from the start of its monetary policy in January 1999. Following the rate-setting meetings of the ECB's decision-making body, the Governing Council, which typically take place on the first Thursday of each month, the ECB announces the monetary policy decisions at 13:45 (CET). Forty-five minutes later, at around 14:30, the ECB President and Vice-President hold a press conference (with the exception of one meeting in summer, where normally no press conference is held). It comprises two elements: a prepared introductory statement that contains the background considerations for the monetary policy decision, and a questions-and-answers (Q&A) part during which the President and the Vice-President are available to answer questions by the attending journalists.

The paper analyzes the ECB's experience from a financial-market perspective. We are interested in knowing to what extent press conferences systematically add relevant information to explain given decisions. The separation of the release of the decision from its explanation allows us to disentangle the effects of monetary policy decisions from those of the accompanying communication. Moreover, because the press conference is broadcasted, and reported upon in real time by financial-market newswire services, it is possible to trace the information flow to financial markets, and thus to separately analyze market reactions to the various types of information.[4] Finally,

[4]This stands in contrast to the information flow for many other central banks, where relevant information on the decisions, such as the minutes of the meetings, is released to the media with an embargo time. In these cases, newswire services prepare a set of news lines that are then released to the markets simultaneously as soon as the embargo time has elapsed. With this simultaneous arrival of news, it is not possible to test the relevance of the various parts of central bank communication.

the Q&A session provides an interesting tool of central bank communication. It gives journalists the opportunity in real time to digest the information provided through the decision and the introductory statement, to compare it with their own prior information, and to ask questions on those issues that need clarification. The analysis in this paper assesses under what circumstances the Q&A session is valuable to clarify issues and the overall message of the press conference.

The main findings of the paper can be summarized as follows. Overall, press conferences have systematically added information. In fact, the size of the market reaction to the press conference is on average substantially larger than the reaction to the policy decision itself, while the press conference at the same time exerts lower effects on market volatility. Moreover, the market reaction to the press conference is related to the characteristics of the decision: the less well a decision has been anticipated by the market, the stronger is the reaction to the introductory statement. This suggests that the statement contains relevant explanations for the reasons underlying the decision, which helps clarify the market participants' interpretation of the decision.

More specifically, the paper asks to what type of information and statements markets react during press conferences. It shows that statements made during the press conference containing a reference to inflationary developments are strong market movers. Furthermore, responses to questions regarding rate discussions at the Governing Council meeting have substantial effects on markets. Other statements—e.g., about the economic outlook, second-round effects, or money growth—are important as well, yet not as consistently as those about inflation and rate discussions.

Finally, the findings on the role of the Q&A session suggest that it does not systematically add information beyond that given in the introductory statement, but it appears to play a clarification role, in the sense that it triggers large financial-market reactions under specific circumstances. In particular, we find that markets are more likely to move in a different direction than in their reaction to the policy decision when there is a high degree of uncertainty among market participants about the state and outlook of the economy. Under situations of elevated macroeconomic uncertainty, the market response to the release of the monetary policy decision itself is muted, suggesting that market participants wait for the clarification provided during the press conference. On the other hand, such directional changes are less likely to occur if the decision itself contains a lot of information (such as when it surprised markets or interest rates were changed).

By looking at financial-market reactions to the announcement of policy decisions and the surrounding communication, this paper is related to different strands of the literature. First, there are numerous studies that analyze market reactions to monetary policy decisions. Most of the work in this literature has focused on the Federal Reserve, though there is increasingly also work on other central banks, including the ECB.[5] This strand of research has reached a consensus that financial-market reactions to the release of monetary policy decisions are substantial.

Second, a number of recent papers analyze issues relating to central bank communication, reflecting the increased importance communication aspects have gained in the conduct of monetary policy over the last decades. Two recent contributions look at the intersection of the announcement of policy decisions and communication, as we do in this study. Gürkaynak, Sack, and Swanson (2005) decompose the policy surprises of Federal Open Market Committee (FOMC) decisions and show that they contain an element of surprise not only about the current decision but also about the future path of interest rates. Given the high degree of predictability of FOMC decisions in recent years, financial markets react predominantly to this "path surprise," which can furthermore be related to the existence of FOMC statements—i.e., communication surrounding the release of the policy decisions. A similar approach has been applied to study the ECB's case in Brand, Buncic, and Turunen (2006), who also find that it is less the announcement of the decision that contains information, but more the press conference that provides substantial new information to financial markets. The present paper shares this finding and goes further by decomposing the elements of the press

[5]Examples of studies on the Federal Reserve are Thornton (1998), Fleming and Remolona (1999), Kuttner (2001), Cochrane and Piazzesi (2002), Bomfim (2003), Ehrmann and Fratzscher (2004), Rigobon and Sack (2004), and Bernanke and Kuttner (2005). Studies covering the ECB are Gaspar, Perez-Quiros, and Sicilia (2001), Hartmann, Manna, and Manzanares (2001), and Ehrmann and Fratzscher (2003). A comparison for the two central banks is provided in Andersson (2008).

conference and by identifying the individual pieces of information to which markets react and which make the press conference constitute a clarifying communication tool.[6]

Another strand of the literature analyzes financial-market reactions to policy decisions and communication, both by committees (Kohn and Sack 2004; Andersson, Dillen, and Sellin 2006; Reeves and Sawicki 2007) and by individual committee members (Reinhart and Sack 2006; Ehrmann and Fratzscher 2007). Research on the role of minutes has emphasized the relevance of timeliness in communication. With the expedited release practices of both the Federal Reserve and the Bank of England, whereby the minutes are now made public prior to the subsequent meeting, financial-market reactions have strengthened considerably (Reinhart and Sack 2006; Reeves and Sawicki 2007). Some (though limited) work has been undertaken on understanding how the media digest information provided by central banks (de Haan, Amtenbrink, and Waller 2004; Berger, Ehrmann, and Fratzscher 2006).[7] Much of this literature analyzes the effect of monetary policy meetings and their announcements; however, to our knowledge the present paper is the first to look in detail—minute by minute and statement by statement—at the individual components of the ECB press conference.

The remainder of this paper is structured as follows. Section 2 provides a simple conceptual framework linking the reaction of financial markets to the information flow of policy decisions and communication. Section 3 presents the data underlying our analysis. Section 4 contains the discussion of the empirical results, together with various extensions and robustness checks. Section 5 focuses on the specific statements contained in the press conferences and analyzes how these have been priced into markets, while section

[6] A number of studies have constructed wording indicators to classify the content of the introductory statements of the ECB's press conferences (Berger, de Haan, and Sturm 2006; Heinemann and Ulrich 2007; Rosa and Verga 2007), showing that there have been significant changes in the tone and the message of these statements—in particular, with regard to the initial years of the ECB—and the effectiveness of certain code words and phrases.

[7] Related studies that focus on the overall role of transparency and communication for different central banks are Guthrie and Wright (2000) and Geraats (2002), or the impact of specific pieces of central bank and other news on financial markets (e.g., Fleming and Remolona 1999; Andersson, Overby, and Sebestyén 2009).

6 specifically investigates to what extent the Q&A part fulfills a clarification role. Finally, section 7 concludes.

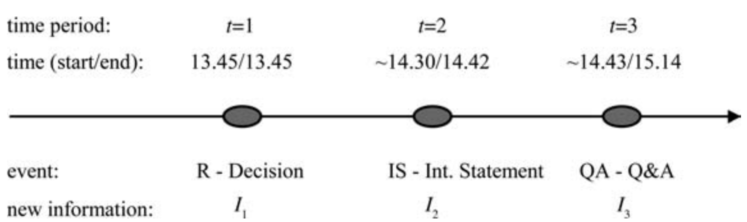
2. The Effect of Press Conferences on Asset Prices: A Simple Framework

This section offers a simple framework for understanding the flow of information on Governing Council meeting days and how this information can affect asset prices. More specifically, we want to sketch how the information that is provided through the three elements of the ECB monetary policy decision—the announcement of the decision, the introductory statement, and the Q&A part—may influence financial markets.

Figure 1 provides an illustration of the sequence with which the ECB provides information to financial markets on meeting days. The first piece of information is the decision itself (I_1), announced at 13:45 CET, which is followed by the introductory statement at 14:30 (I_2), and then finally the subsequent Q&A part of the press conference (I_3).

While these events occur at different points in time, it is important to note that they are not independent. To give an example, let us assume the Governing Council makes a decision that is somewhat unexpected by financial markets. This will imply that the news content in the release of the decision is relatively large. At the same time, knowing that this decision had not been perfectly predicted, the ECB may wish to provide an explanation for the surprising decision. More often than not, this will imply that in the introductory statement as well, the news content is relatively large. Contrasting

Figure 1. Timeline of ECB Monetary Policy Decision and Communication



this with a case where the decision is as expected—i.e., contains no news, such that the introductory statement most likely reaffirms the views of market participants as well—shows that one cannot exclude the possibility that the news content of the introductory statement depends on the news contained in the release of the decision. This endogeneity becomes even clearer in the case of the Q&A session, where the questions obviously are responsive to the earlier pieces of news. This can be formulated as follows: $I_2 = \phi_2(I_1)$ and $I_3 = \phi_3(I_1, I_2)$.

To understand how financial markets will respond to this flow of information, it is useful to think of the price of a financial asset p_t in a present-value setting, whereby it reflects agents' expectations of future fundamentals f_{t+i}: $p_t = (1-\theta)\sum_{i=0}^{\infty}\theta^i E_t(f_{t+i}|\Omega_t)$, with Ω_t as the information set at time t, E as the expectations operator, and θ as the discount factor. Given the short time window we are interested in during the announcement and press conference, we ignore the discount factor. By defining a vector f that contains the fundamentals of all future periods f_{t+i}, we can formulate a dynamic expression of the change of the asset price Δp_{t+1} as

$$\Delta p_{t+1} = E_{t+1}(f|\Omega_t, I_{t+1}) - E_t(f|\Omega_t) = \Delta E_{t+1}(f|\Omega_t, I_{t+1}), \quad (1)$$

in which the emergence of new information I_t is assumed to be the only factor inducing a price change at time $t+1$. Note that we assume that the vector of fundamentals does not change through the arrival of news—it is merely agents' expectations that are affected. This simplification assumption is also made on the basis of the short time window we are interested in.

In this framework, the response of asset prices to news depends on the way agents update their expectations about the underlying fundamental f. In the simplest case, this process can be described by a linear relationship,

$$\Delta p_{t+1} = \Delta E_{t+1}(f|\Omega_t, I_{t+1}) = \beta\ I_{t+1}, \quad (2)$$

the assumption usually applied in the literature on announcement effects.[8] Taking into account the three-tiered information flow on

[8]A similar model framework is provided in Faust et al. (2007).

the occasion of Governing Council meetings, this amounts to the following relationship:

$$\Delta p_{1,R} = \beta_1\, I_1 + \varepsilon_1 \tag{3a}$$

$$\Delta p_{2,IS} = \beta_2\, I_2(I_1, \varepsilon_1) + \varepsilon_2 \tag{3b}$$

$$\Delta p_{3,QA} = \beta_3\, I_3(I_2, \varepsilon_2, I_1, \varepsilon_1) + \varepsilon_3. \tag{3c}$$

Equations (3a)–(3c) take into account the endogeneity of the information flow. Moreover, since we can measure the information I_t only imperfectly, we also include the unobserved information components ε_t in each step.

While simple, it might not be very realistic to assume that the updating of expectations is a linear function of the news. For instance, in the presence of a rather noisy information set, the implications of a news item on fundamentals might be less clear-cut, such that there is only a muted response to news. In that case, a more generalized setup can be imagined whereby the updating depends on the information set currently available.

The aim of this stylized framework is to illustrate the information flow and the updating of agents' expectations, as well as the corresponding movements of asset prices and trading activity. The next step is to identify the information components in the various steps of the ECB decision and its communication, to which we turn now.

3. Data

This section discusses the main data used in the empirical analysis—foremost, the three-month Euribor futures rates, the newswire, and other data on ECB press conferences, as well as the proxies for macroeconomic uncertainty.

3.1 Three-Month Euribor Futures

This paper analyzes the reaction of three-month Euribor futures to the communication on Governing Council meeting days, given the fact that this is the most traded money-market instrument on this occasion. We have obtained intraday data from Tick Data, Inc. The prices are recorded as actual transaction prices on LIFFE on a tick-by-tick basis.[9] Because these observations are unequally spaced, we calculate price data on a minute-by-minute frequency by linear interpolation of the two tick prices immediately before and after the full minute (Andersen et al. 2003). For an analysis of trading activity, we furthermore obtain the number of ticks recorded within a given minute. In addition, although only as of July 2003, the data contains information on traded volumes, measured as the number of contracts (over €1 million each) traded.

The decision to calculate minute-by-minute data arises because this is the frequency at which we can obtain data on the news headlines by the financial newswires (described below). From the price data, we calculate returns as $r_t = 100 * [\ln(p_t) - \ln(p_{t-1})]$. An alternative measure for the market evolution would consist in the first difference of prices, as the implied futures rate i_t is derived from the quoted price by subtracting the latter from 100, such that $i_t - i_{t-1} = (100 - p_t) - (100 - p_{t-1}) = p_{t-1} - p_t$. The two measures are extremely similar, with a 1 percent return being roughly equivalent to a 100-basis-point decrease in the implied futures rate. Finally, we construct a measure of realized volatility based on Andersen et al. (2003) as the sum of the squared returns over the relevant time windows.[10]

As is well known, such high-frequency financial-market data are subject to intraday patterns and day-of-the-week effects, which will have to be controlled for in any subsequent analysis.

[9]Euribor futures contracts are based on an interbank rate, which is highly correlated with the ECB's policy rate. The data generally refer to the contract with the nearest maturity. The switch to the next maturity is done by a procedure that compares daily tick volumes for two adjacent contracts. It switches usually around three to five days before expiration of the contract with the nearest maturity, when daily tick volumes exceed those of the old contract. This procedure ensures maximum liquidity of the considered contracts. For more information, see http://www.tickdata.com.

[10]Choosing a length of the time window over which realized volatility is calculated, and the frequency of the underlying return data, is subject to a trade-off (Andersen et al. 2003). In our case, the minutely frequency of the return data is naturally given by the frequency of some of the explanatory variables. The time window over which we calculate realized volatility similarly arises naturally, through the length of the various parts of the press conference.

3.2 Monetary Policy Decisions and the Press Conference

Information on the ECB's monetary policy decisions and press conferences has been obtained from its web site. The taped versions of the press conferences on Bloomberg allow us to determine the length of the introductory statement and the Q&A session, respectively, for each press conference. Due to data availability, our sample starts with the press conference in July 2001; it ends with the conference in April 2006, such that our sample contains fifty-three observations. It is therefore important to keep in mind that results are based on a small sample. Table 1 provides a few summary statistics for the press conferences in our sample. It has lasted on average around forty-four minutes, with twelve minutes taken up by the reading of the introductory statement and thirty-two minutes for the Q&A session. On average, there are around sixteen questions asked in the Q&A session. However, all figures vary substantially over time. The number of questions posed, for instance, varies from eight on August 30, 2001, to thirty-one on June 5, 2003 (interestingly, both days on which policy rates were changed).

As we are inter alia interested in market reactions to individual statements made during the press conference, we extract the real-time reports (snaps) released on a commonly used newswire service, Reuters News. As the snaps are available from Reuters for thirteen consecutive months only, our sample starts only in September 2004.[11] Furthermore, the sample ends in July 2005 (note that

Table 1. Summary Statistics for the ECB's Press Conference

	Average	Minimum	Maximum
Length of Press Conference	43.77	26	72
Length of Introductory Statement	11.92	8	19
Length of Q&A Session	31.85	16	54
Number of Questions during Q&A Session	16.36	8	31

Note: Statistics based on fifty-three press conferences from July 2001 to April 2006.

[11] Alternative sources like Bloomberg or Market News International provide these data for considerably shorter periods only.

no press conferences are held in August), in order to restrict the analysis to a relatively homogeneous time sample—namely, a period where markets did not expect any immediate changes in policy rates.

As an illustration, table 7 in the appendix provides the snaps released on Reuters during the press conference in November 2004. Each snap consists of a brief statement, reporting about the main points made during the press conference. Importantly for our purposes, the time stamp is available for each snap, such that we know the exact minute at which the information reaches the markets. We distinguish the snaps according to their content, differentiating between statements on the economic outlook, inflation, second-round effects, money growth, and interest rates.[12] The latter classification was chosen for statements that relate directly to the discussion on policy rates in the Governing Council. Such statements are never made during the introductory statement but are sometimes made in response to a question (such as whether a rate decision was made unanimously or whether the Governing Council has discussed all options—i.e., increasing, decreasing, maintaining interest rates, etc.) during the Q&A session. From the snaps, we construct a time series for each of the content categories, which is equal to one in any minute where an according snap is recorded on Reuters, and equal to zero otherwise.

A number of caveats of this methodology should be emphasized. First, newswire services may wrongly report or misinterpret a statement. However, as our objective is to assess communication from the perspective of financial markets, it is important to analyze the information market participants actually receive. Second, there are a number of newswire services that report in real time, and the press conference is furthermore televised. Accordingly, financial-market

[12]Our data set contains 530 snaps. Of these, 483 have an economic content (as opposed to snaps reporting that the ECB President opens the press conference or the Q&A session, or snaps related to topical issues other than monetary policy or the economic developments, such as central bank gold sales). Our classification covers two-thirds of the statements with economic content. Snaps not covered relate, for instance, to global imbalances or fiscal policy. Their inclusion does not alter the results of our econometric analyses.

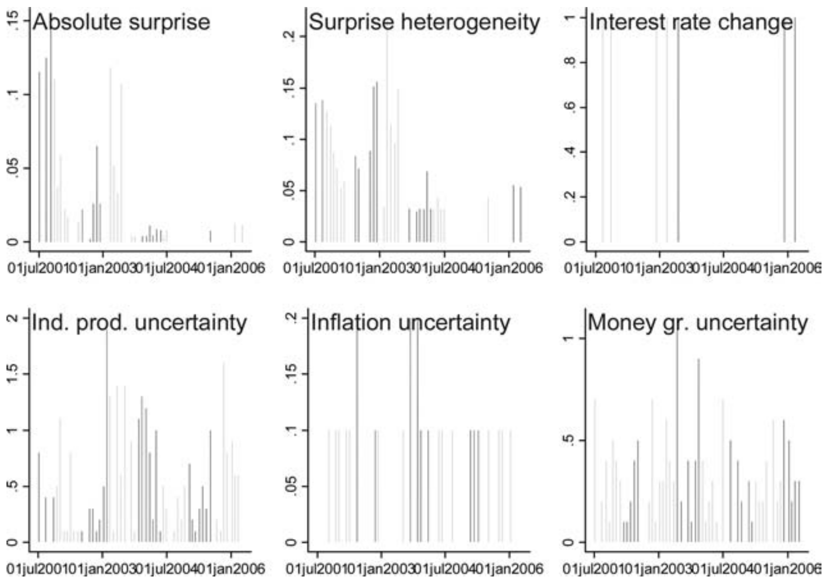
participants might receive different information, depending on their source. However, a comparison of the snaps released by Reuters, Bloomberg, and Market News International shows no major differences with respect to their timing and content. Furthermore, the delay in newswire reports relative to the televised version is minimal. Finally, to ensure that we are measuring the effect of the ECB's communication, rather than other news, we control for the market reaction to the release of U.S. jobless claims figures, which occurs at 14:30 on Thursdays. We do so by calculating the surprise component contained in the released figures as the actual release minus market expectations measured through the median response of a Bloomberg survey among market participants.

Finally, we are interested in obtaining measures that characterize a given policy decision. First, we obtain information on the decision from the ECB web site, and define a dummy variable that is equal to one when interest rates have been changed, and to zero otherwise. Furthermore, for a measure of the surprise component contained in a decision, we employ the results of a Reuters survey among market participants, which is conducted a few days prior to the Governing Council meeting. The surprise component in the decision is constructed as the difference between the decision and the mean response in the survey. Of interest in our analysis is the absolute value of this surprise component. The second proxy for the surprise relates to the heterogeneity in expectations across market participants. For that purpose, we calculate the standard deviation of expectations across individual analysts participating in the Reuters survey. As shown in figure 2, this measure of heterogeneity in market expectations is highly positively correlated with the absolute surprise. In order to obtain uncorrelated regressors for our econometric analyses, we obtain the residuals of a regression of the absolute surprise on the heterogeneity measure, estimated in a simple OLS regression.

3.3 Macroeconomic Uncertainty

The final type of data used in this paper (also shown in figure 2) relates to macroeconomic uncertainty, as we are interested in the effects of the press conference conditional on the macroeconomic

Figure 2. Data on Characteristics of Monetary Policy Decisions and Macroeconomic Uncertainty



Note: The figure shows data on characteristics of monetary policy decisions and macroeconomic uncertainty, with the latter measured as the absolute difference between the latest macroeconomic release prior to a Governing Council meeting and the corresponding market consensus (derived as the median response of a Bloomberg poll among financial-market analysts a few days prior to the release). Sample period: July 2001 to April 2006.

environment.[13] However, macroeconomic uncertainty is obviously hard to measure. Our proxy makes use of the surprise component in macroeconomic releases for euro-area industrial production, HICP inflation, and money growth by subtracting the announced figures from market expectations (as measured by the median response in corresponding Bloomberg surveys). For each of these variables,

[13]Gropp and Kadareja (2006) show that stock market reactions to news depend on the quality of public information. With lower-quality public information, the stock market reacts with more volatility to news, suggesting that better public information lowers the extent to which traders differ in their interpretation of new information. In a similar vein, we might expect that the market response to a monetary policy decision is affected by the degree of macroeconomic uncertainty.

we obtain the latest release that occurred prior to a Governing Council meeting and use the surprise component contained therein as our measure of macroeconomic uncertainty at this point in time.

4. The Effect of the ECB's Meeting-Day Communication

We start by estimating the relevance of the ECB's meeting-day communication by comparing market developments on meeting days to nonmeeting days (section 4.1) before turning to the specific market reactions to the individual components of ECB decisions and communications (section 4.2).

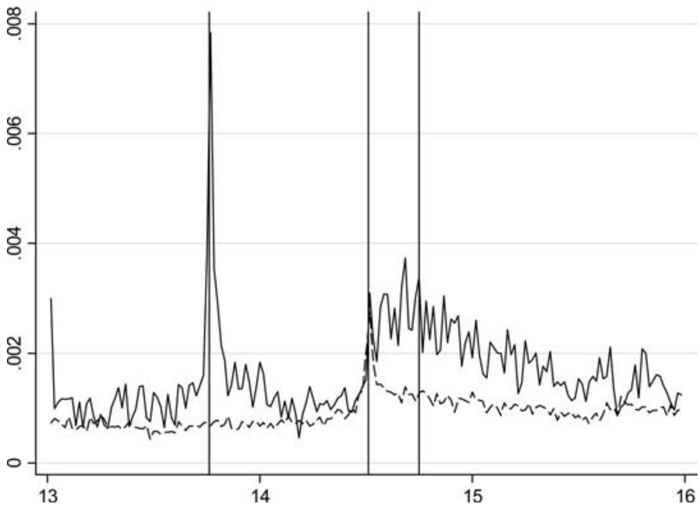
4.1 Relevance for Financial Markets

Figures 3–5 illustrate how the three-month Euribor futures market behaves on the days of the ECB's Governing Council meetings. For each minute from 13:00 to 16:00, the solid lines show the average absolute return (figure 3), the average number of ticks (figure 4), and the average volume traded (figure 5) on days of ECB Governing Council meetings and press conferences. For a comparison, the same statistics, measured on Thursdays without Governing Council meetings, are shown by the dashed line.[14]

A number of interesting facts are apparent from the figures. First, there are clear intraday patterns in market behavior. On both ECB meeting days and other Thursdays, market activity picks up considerably in the afternoon, which coincides with the opening of the U.S. markets. In particular, the weekly release of U.S. jobless claims at 14:30 leads to a spike in absolute returns, ticks per minute, and traded volume alike. Second, the effects of the release of the monetary policy decision at 13:45 and of the press conference, which starts at around 14:30, are also clearly discernible. Market activity rises considerably at 13:45 and remains elevated for a considerable period of time. Just before the start of the press conference, market

[14]Days with a Governing Council meeting but without a press conference are excluded from the calculation of both lines shown in the figures. The comparison group is calculated for Thursdays exclusively in order to avoid the possibility that day-of-the-week patterns in financial-market behavior affect their properties.

Figure 3. Average Absolute Returns on Press Conference Days versus Thursdays without Governing Council Meetings

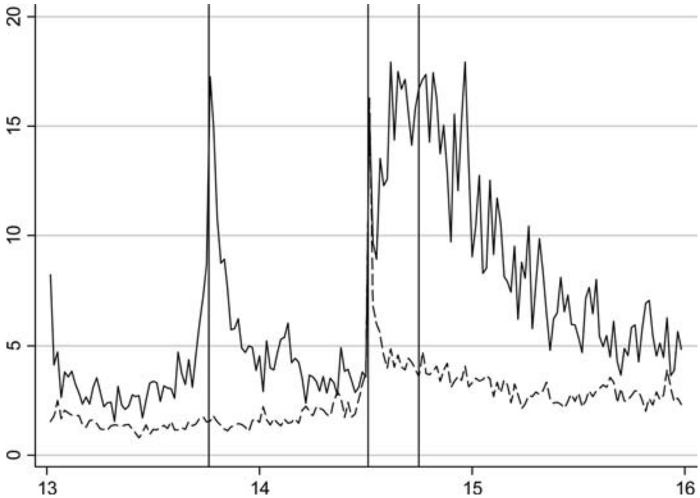


Note: The figure shows average absolute returns per minute in three-month Euribor futures, on days with a press conference (solid line) versus benchmark Thursdays without Governing Council meetings (dashed line), for a time window from 13:00 to 16:00. Vertical lines show the time of the release of the decision, the start of the press conference, and the average start of the Q&A session. Sample period: July 2001 to April 2006.

activity is roughly back to normal. The effects of the press conference appear in the data a couple of minutes after 14:30. This is to be expected, not only because the press conference sometimes starts with a slight delay, but also because it does not immediately start with information to which a market reaction should be expected: the ECB President first welcomes all participants, often informs about the attendance at the Governing Council meeting (e.g., if the president of the Ecofin has attended), and starts by reiterating what decision has been made at the meeting, which is of course known to markets since 13:45.

Beyond this graphical inspection, table 2 reports the outcome of some statistical tests. Absolute returns, ticks per minute, realized

Figure 4. Average Number of Ticks per Minute on Press Conference Days versus Thursdays without Governing Council Meetings

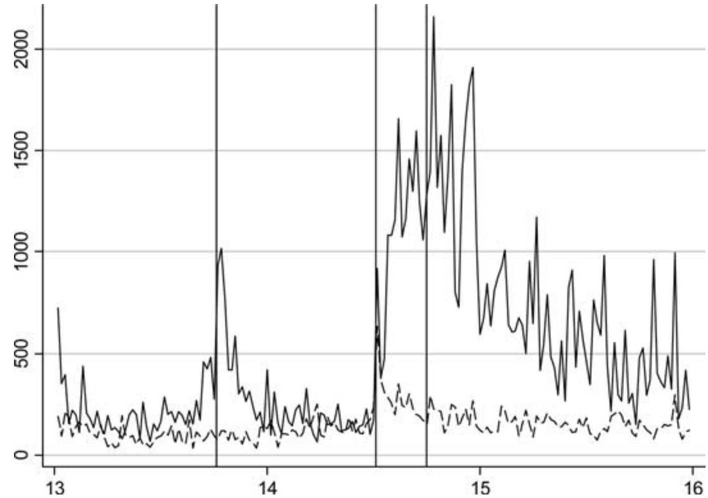


Note: The figure shows average ticks per minute in three-month Euribor futures, on days with a press conference (solid line) versus benchmark Thursdays without Governing Council meetings (dashed line), for a time window from 13:00 to 16:00. Vertical lines show the time of the release of the decision, the start of the press conference, and the average start of the Q&A session. Sample period: July 2001 to April 2006.

volatility (calculated per minute, as the length of time windows differs), and volume are compared for different time windows on press conference days and on benchmark Thursdays without Governing Council meetings through simple mean comparison tests. The first column compares market reactions to the release of the monetary policy decision in a ten-minute window—i.e., from 13:45 through 13:54—with market developments in the control window on non-meeting days.[15] The second column compares the market activity during the reading of the introductory statement—based on averages for starting time and length, as recorded on Bloomberg, namely from

[15] All results related to the effect of the release of the decision in this paper will be based on this ten-minute window; none of the results is affected significantly when extending this time window.

Figure 5. Average Volume Traded per Minute on Press Conference Days versus Thursdays without Governing Council Meetings



Note: The figure shows the average volume traded per minute in three-month Euribor futures, on days with a press conference (solid line) versus benchmark Thursdays without Governing Council meetings (dashed line), for a time window from 13:00 to 16:00. Vertical lines show the time of the release of the decision, the start of the press conference, and the average start of the Q&A session. Sample period: July 2001 to April 2006.

14:32 through 14:43—with a control window on nonmeeting days. The third column provides estimates of the effect of the Q&A session. As the length of the Q&A sessions varies substantially (see table 1), often covering various topics unrelated to monetary policy toward the end, we decided to cut off the analysis after fifteen minutes. Such an approach also seems justified by the financial newswire coverage of the press conference: snaps typically become less frequent toward the end of the press conference. Finally, the fourth column shows market reactions for the combined introductory statement and Q&A session.

All four tests—for returns, tick numbers, volatility, and volume—clearly show evidence for substantially increased market activity on meeting days, with all differences being significant at the 99 percent

Table 2. Market Effects of the ECB's Press Conference

	Release of Decision			**Intr. Statement**			**Q&A**			**Entire Press Conference**		
	Actual	Control	Diff.	Actual	Control	Diff.	Actual	Control	Diff.	Actual	Control	Diff.
Absolute Return	0.006	0.002	***	0.012	0.004	***	0.010	0.004	***	0.018	0.006	***
Ticks per Minute	9.051	1.413	***	14.980	4.806	***	14.173	3.755	***	14.522	4.222	***
Realized Volatility	0.329	0.021	***	0.191	0.055	***	0.161	0.047	***	0.173	0.051	***
Volume per Minute	513.422	92.014	***	1210.451	265.961	***	1308.326	198.638	***	1278.749	228.559	***

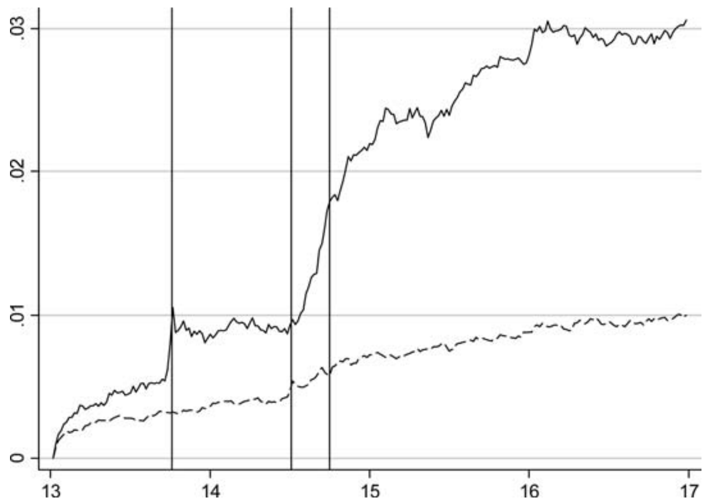
Note: All figures are calculated for the actual time windows of ECB communication ("Actual") and a corresponding control time window on non-announcement days ("Control"). "Difference" represents results for tests of equality. "Release of the Decision" relates to the ten minutes following the release of the ECB's monetary policy decisions, and a time window from 13:45 through 13:54 on non-announcement Thursdays. Figures calculated for the various parts of the press conference are compared with a time window on other Thursdays, from 14:32 through 14:43 for the introductory statement, and from 14:44 through 15:58 for the Q&A session. *, **, and *** denote significance at the 90 percent, 95 percent, and 99 percent level, respectively.

level (as indicated by the stars in the column labeled "Diff."). Moreover, an important stylized fact is that the market reaction to the entire press conference is substantially higher than the market reaction to the announcement of monetary policy decisions. On average, the absolute return reaction to the whole press conference is about three times stronger than the market reaction to the announcement of the policy decision. The figures in column 1 are significantly larger than those in column 4 at the 1 percent level for absolute returns and ticks per minute, and at the 5 percent level for volume. While being an important market mover, it is striking that the effect of the press conference is digested by financial markets in a relatively smooth fashion. A comparison of the realized volatility measures shows that the large effect of the press conference occurs with only half of the volatility compared with the release of the decision (statistically significantly smaller at the 6 percent level). These results underscore the importance of the press conference as a central source of information.

To assess the news content contained in the ECB's communication on Governing Council meeting days, figure 6 shows the price movements in three-month Euribor futures on meeting and nonmeeting days. Prices are normalized at 0 at 13:00, and their absolute level relative to the price at 13:00 is plotted for the subsequent four hours. There is clear evidence that prices move by substantially more on meeting days, and that the movements are highly persistent. A jump in prices is observed on the occasion of the release of the decision; prices remain basically flat thereafter until the beginning of the press conference, then continue moving in the same direction of the initial jump until around 16:00, after which they once more remain basically flat.

There is therefore clear evidence that the release of monetary policy decisions and the ensuing press conference are considered relevant by financial-market participants, and that substantial amounts of news are priced into the markets on Governing Council meeting days. Furthermore, the magnitude of the effect is sizable. The average absolute return, for instance, rises by a factor of around 3 relative to Thursdays without Governing Council meetings. The volume of trade increases by even more—both the number of ticks per minute and volume increase by a factor of around 6 during the release

Figure 6. Price Movements in Three-Month Euribor Futures Contracts on Press Conference Days versus Thursdays without Governing Council Meetings



Note: The figure shows the average evolution of prices of three-month Euribor futures compared with the level of prices pertaining at 13:00, in absolute terms, on days with a press conference (solid line) versus benchmark Thursdays without Governing Council meetings (dashed line), for a time window from 13:00 to 17:00. Vertical lines show the time of the release of the decision, the start of the press conference, and the average start of the Q&A session. Sample period: July 2001 to April 2006.

of the monetary policy decision compared with non-announcement days. Moreover, press conferences appear to be a substantial market mover—even more so than the announcements of monetary policy decisions themselves—while at the same time leading to relatively little market volatility given the magnitude of the observed market moves.

4.2 Determinants of Market Reactions

Having seen that markets react strongly to the ECB's communication, we want to understand better what factors determine market reactions to the different communication events on Governing Council meeting days. In the search for these determinants, we attempt to explain the absolute returns,[16] ticks per minute, and market volatility as observed on the fifty-three meeting days by a number of factors, in a regression model of the type

$$y_{R,t} = \alpha_{1,R} + \sum_i \beta_{i,R} x_{i,t} + \sum_j \gamma_{j,R} c_{j,t} + \varepsilon_{R,t} \tag{4a}$$

$$y_{IS,t} = \alpha_{1,IS} + \sum_i \beta_{i,IS} x_{i,t} + \sum_j \gamma_{j,IS} c_{j,t} + \delta_{1,IS} \hat{\varepsilon}_{R,t} + \varepsilon_{IS,t} \tag{4b}$$

$$\begin{aligned} y_{QA,t} = {} & \alpha_{1,QA} + \sum_i \beta_{i,QA} x_{i,t} + \sum_j \gamma_{j,QA} c_{j,t} + \delta_{1,QA} \hat{\varepsilon}_{R,t} \\ & + \delta_{2,QA} \hat{\varepsilon}_{IS,t} + \varepsilon_{QA,t}, \end{aligned} \tag{4c}$$

where y is average absolute returns, average ticks per minute, or market volatility, as measured over the relevant time windows for the release of the decision (y_R, 13:45–13:54, equation (4a)), for the introductory statement (y_{IS}, equation (4b)), and for the Q&A session (y_{QA}, equation (4c)), respectively. t denotes the day of a Governing Council meeting, such that $t = 1, \ldots, 53$. When modeling the average number of ticks per minute, we include a time trend to allow for increasing market depth for this variable (which does not enter significantly in the other models and is therefore not included elsewhere).

The models contain three types of explanatory variables. First, we include proxies for the informational content—akin to I_1 in the simple framework of section 2—of the release of the decision, summarized in the terms $\sum_i \beta_i x_{i,t}$, which relate to the characteristics of the decision itself. Second, as discussed in section 2, we allow for an effect of the unexplained parts of the market reaction to the policy decision $\hat{\varepsilon}_{R,t}$ on the market behavior during the introductory statement and Q&A part, as well as for an effect of the unexplained part of the introductory statement $\hat{\varepsilon}_{IS,t}$ on the market during the Q&A session. Again, we see these as proxies for the informational content of the various announcements.

[16]Note that we are analyzing the response of the *absolute* returns, rather than returns, mainly because some of the explanatory variables are unsigned by definition.

Finally, we control for the degree of market uncertainty before the decision, proxied by realized volatility from 10:00–13:00 in the morning of Governing Council meeting days and by the degree of macroeconomic uncertainty. These variables are denoted as controls $c_{j,t}$.

The regression results are reported in table 3. Turning first to the *characteristics* of the policy decisions, it is clear that markets react more to the release of the decision the larger the surprise component in a given decision (first set of results in the three panels of table 3). This has to be expected, as a more surprising decision contains more news, and this requires a stronger market adjustment. Heterogeneity in market expectations does not appear to exert any effect on absolute returns, whereas it affects the number of trades exercised after the release of the decision. Heterogeneity in expectations around a certain mean (which the model has controlled for by means of the surprise component contained in the decision) does not increase the *average* news component of the decision; it only implies that more individual market participants will have to rebalance their market positions. The results are in line with this reasoning. Finally, market activity is furthermore rising in the case of changing policy rates, without any further effect on absolute returns or market volatility. The effects are relatively sizable. A one-standard-deviation increase in the size of the absolute surprise leads to 4.7 more ticks per minute (an increase of more than 50 percent of the average 9 ticks per minute recorded in table 2); a one-standard-deviation increase in market heterogeneity leads to around 4 extra ticks (a 44 percent increase).

While several of the explanatory variables are significantly estimated in model (4a)—i.e., for the market reaction to the release of the decision—only a few explanatory factors emerge for the press conference, i.e., the introductory statements and the Q&A sessions (see second and third set of results in table 3). In particular, average absolute returns and market volatility during the introductory statement depend on the magnitude of the surprise component contained in a monetary policy decision—the larger this component, the bigger is the market reaction during the introductory statement. This is in line with the hypothesis that the news content of the introductory statement is endogenous with regard to the news content of the decision—in case of a surprise, explanations are provided to the

Table 3. Determinants of the Market Effects of the ECB's Press Conference

	Release of Decision				Introductory Statement				Q&A			
	Actual		Control		Actual		Control		Actual		Control	
Absolute Return	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**
Proxies for News Content												
Characteristics of the Decision												
Absolute Surprise	0.095**	0.036	—		0.081***	0.027	—		0.092	0.056	—	
Surprise Heterogeneity	−0.005	0.033	—		0.037	0.037	—		0.052	0.062	—	
Interest Rate Change	0.001	0.004	—		−0.003	0.004	—		−0.009	0.006	—	
Residuals												
Release of Decision	—		—		0.786***	0.220	−0.088	0.160	−0.063	0.165	0.331**	0.132
Introductory Statement	—		—		—		—		0.284**	0.112	0.126**	0.057
Proxies for Uncertainty												
Prior Realized Volatility	0.019	0.048	0.017	0.012	0.023	0.044	0.071	0.044	0.034	0.034	0.085***	0.027
Macro Uncertainty												
Industrial Production	−0.001	0.002	0.000	0.000	0.002	0.003	0.001	0.001	0.000	0.004	−0.001	0.001
Inflation	−0.040***	0.011	0.003	0.003	0.049**	0.024	0.005	0.007	−0.048	0.032	−0.009	0.006
Money Growth	0.001	0.005	−0.001	0.001	0.004	0.005	−0.003	0.002	−0.008	0.005	0.001	0.001
# of Observations	53		188		53		188		53		188	
R-Squared	0.398		0.024		0.350		0.069		0.167		0.057	
Control: Time Trend	No		No		No		No		No		No	

(continued)

Table 3. (Continued)

	Release of Decision				Introductory Statement				Q&A			
	Actual		Control		Actual		Control		Actual		Control	
Ticks per Minute	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**
Proxies for News Content												
Characteristics of the Decision												
Absolute Surprise	116.870***	23.499	—		49.160	44.039	—		110.825	79.616	—	
Surprise Heterogeneity	132.662***	28.280	—		25.986	39.142	—		50.044	71.901	—	
Interest Rate Change	6.803***	2.293	—		6.257	5.674	—		−0.265	6.438	—	
Residuals												
Release of Decision	—		—		0.969**	0.362	0.222	0.144	0.875***	0.224	0.265*	0.145
Introductory Statement	—		—		—		—		0.592***	0.145	0.367***	0.084
Proxies for Uncertainty												
Prior Realized Volatility	−11.759	20.197	11.069	6.838	−33.676	59.239	60.090***	17.640	−36.721	53.307	55.628***	11.553
Macro Uncertainty												
Industrial Production	0.948	1.377	−0.140	0.237	−0.942	2.521	0.842	0.744	−2.472	3.168	0.155	0.473
Inflation	−30.231***	9.174	0.643	2.763	22.502	21.777	4.432	6.055	−34.217	21.894	−0.880	3.899
Money Growth	−5.137*	2.786	0.452	0.432	4.072	5.710	−1.550	1.402	0.989	5.415	−0.181	0.844
# of Observations	53		188		53		188		53		188	
R-Squared	0.764		0.022		0.370		0.086		0.526		0.323	
Control: Time Trend	Yes		Yes		Yes		Yes		Yes		Yes	

(continued)

Table 3. (Continued)

	Release of Decision				Introductory Statement				Q&A			
	Actual		Control		Actual		Control		Actual		Control	
Realized Volatility	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**
Proxies for News Content												
Characteristics of the Decision												
Absolute Surprise	12.203***	3.772	—		1.291*	0.728	—		0.115	0.818	—	
Surprise Heterogeneity	1.233	3.329	—		1.375	0.932	—		1.035	1.147	—	
Interest Rate Change	0.477	0.376	—		−0.050	0.094	—		−0.028	0.103	—	
Residuals												
Release of Decision	—		—		0.175*	0.101	0.261	0.226	−0.001	0.047	0.198	0.162
Introductory Statement	—		—		—		—		0.269*	0.154	0.282***	0.076
Proxies for Uncertainty												
Prior Realized Volatility	−0.668	1.292	0.319***	0.101	−0.160	0.879	1.395***	0.421	0.714	1.204	1.527**	0.694
Macro Uncertainty												
Industrial Production	0.031	0.159	−0.006*	0.003	0.009	0.065	−0.001	0.007	−0.080	0.078	−0.014	0.009
Inflation	0.213	0.722	0.015	0.030	0.340	0.431	−0.028	0.075	−0.582	0.538	−0.012	0.091
Money Growth	−0.690	0.484	0.009	0.007	0.193	0.167	−0.013	0.023	0.021	0.091	−0.018	0.027
# of Observations	53		188		53		188		53		188	
R-Squared	0.542		0.091		0.344		0.200		0.093		0.107	
Control: Time Trend	No		No		No		No		No		No	

Notes: Results are based on estimation of model (4a)–(4c), using three-month Euribor futures. *, **, and *** denote significance at the 90 percent, 95 percent, and 99 percent level, respectively.

public for the reasons behind the given decision. Interestingly, there is no further relationship between the size of the surprise and market reactions during the Q&A session, which could indicate that the explanations in the introductory statement have provided sufficient information to the public, such that no further need for clarification in that respect arises during the Q&A session.

Furthermore, returns, trading activity, and realized volatility are mostly significantly related to the unexplained component of the release of the decisions—i.e., the residuals $\hat{\varepsilon}_{R,t}$—suggesting that large market moves in reaction to the release of the monetary policy decision are generally also followed by large moves during the press conference. Again, this is in line with the idea that lots of news in one part of the event chain is followed by lots of news in the other parts.

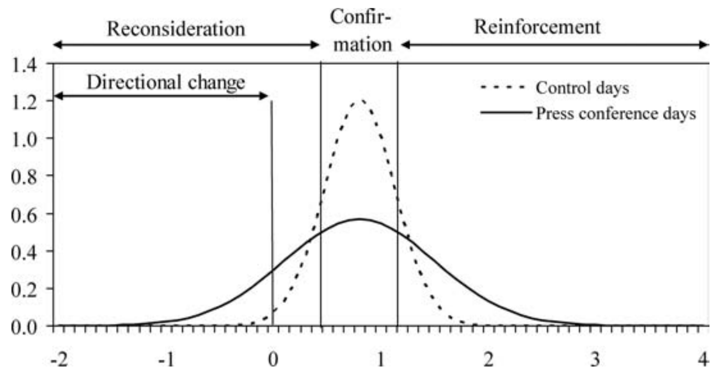
Finally, table 3 shows that while prior market volatility does not affect the market reactions to ECB decisions and to press conferences, the degree of macroeconomic uncertainty does. With increasing uncertainty about inflation developments in particular, the market reaction to the release of monetary policy decisions becomes muted, a pattern not observed on days without press conferences. The effects imply that a one-standard-deviation increase in the size of uncertainty about inflation leads, e.g., to a reduction by 1.8 ticks (roughly 20 percent of the average number of ticks recorded).

5. The Clarification Objective of the Q&A Session

An interesting feature of the ECB's press conference is its Q&A session, which provides journalists with an opportunity to ask clarification questions, whereby the news content is clearly endogenous relative to the previous communication events. This section analyzes whether there is indeed evidence for such a clarification role, and under what conditions.

Our empirical approach is based on the following considerations. In the absence of a counterfactual—i.e., an estimate of how financial markets would have evolved after the reading of the introductory statement, but without a subsequent Q&A session—we assume that market developments tend to be persistent, as it takes time until the arrival of earlier information (in our case, the information provided through the introductory statement) is correctly priced (see,

Figure 7. Stylized Description of Market Movements with and without Arrival of New Information



Note: The figure plots hypothetical distributions of the autoregressive coefficient of market movements under two scenarios: (i) the absence of new information (dotted line) and (ii) the arrival of new information (solid line). In both cases, the mean coefficient stands at 0.8, indicating the persistence of market movements across subsequent time windows. With the arrival of new information, market developments are more likely to deviate from this coefficient. Market movements in the vicinity of 0.8 suggest that the new information has roughly confirmed previous information. Substantially larger coefficients point to reinforcing information, as market movements continue in the same direction, yet are strengthened. Substantially smaller coefficients arise if the new information corrects earlier information, leading to a weakening of market moves or even directional changes (with autoregressive coefficients below 0).

e.g., Evans and Lyons 2005). Once new information arrives (in our case, the information contained in the Q&A session), earlier market moves can either be confirmed, reinforced, or reconsidered. For the latter case, the trend movements can either weaken while continuing in the same direction or they can change direction. These possibilities are depicted in a stylized fashion in figure 7. As is clear from the figure, both "reinforcement" and "reconsideration" become more likely with the arrival of new information. For testing purposes, the relevant question is where to locate the dividing lines between "confirmation," "reinforcement," and "reconsideration." In the absence of a clear prior on the location of these lines, the most objective criterion is the dividing line between directional changes and continuations of the direction of earlier market moves. Our hypothesis is therefore that a clarification role of the Q&A session should lead to more frequent directional changes.

As a starting point, table 2 shows that, on average, market movements initiated during the introductory statement are continued during the Q&A session, as the absolute return measured over the entire press conference is substantially larger than during the introductory statement (0.018 versus 0.012). A mean comparison test shows that this difference is statistically significant, with a p-value of 0.05. However, at the same time, it is also apparent that there are instances where the market movement during the introductory statement does change direction, as the sum of the absolute return during the introductory statement and during the Q&A session (0.012 and 0.010) add up to more than the absolute return during the entire press conference (0.018). Importantly, the equivalent test for non-press-conference days gives a different picture, where absolute returns during the control window for the entire press conference are bigger than during the control window for the introductory statement with a p-value of 0.01. This suggests that there is less variance on non-press-conference days, or in other words that there are fewer cases of a directional change.

Table 4 provides a more direct comparison. It calculates the relative share of directional changes. On control days, this occurs consistently in less than 50 percent of all cases, suggesting that market movements are indeed somewhat persistent. By comparison, on press conference days, directional changes are more likely. In 60 percent of all cases, the market move following the release of the decision tends in the opposite direction than the move during the Q&A session, which is significantly larger than the corresponding number on control days (namely 44 percent), at the 95 percent significance level.[17]

Taken together, this evidence is in line with the hypothesis that the Q&A session fulfills a clarification role, as (i) the *size* of market movements is significantly larger during Q&A sessions than during comparable times and (ii) the *direction* of movements is significantly different, as the higher likelihood of directional changes indicates.

As the next step, we want to know under which circumstances this clarification objective is particularly useful. In other words, we

[17]Note, however, that this number is insignificantly different from 50 percent.

Table 4. Probability of Directional Changes

	Actual	Control	Difference (p-value)
Introductory Statement vs. Release of Mon. Policy Decision	0.472	0.466	0.473
Q&A vs. Release of Mon. Policy Decision	0.604	0.444	0.020
Q&A vs. Introductory Statement	0.491	0.406	0.140

Note: All figures are calculated for the actual time windows of ECB communication ("Actual") and a corresponding control time window on non-announcement days ("Control"). "Difference" represents the p-value of the test of equality.

would like to identify the determinants of directional changes. To conduct such an analysis, we create a discrete dummy variable that is equal to one in the case of a directional change. We model this variable (for which we have fifty-three observations) by means of a probit specification, containing the same regressors as model (4a)–(4c) above.

Table 5 provides the corresponding results. Positive parameters raise the probability that the dependent variable equals one, i.e., that a directional change has occurred. The table reports marginal effects, i.e., the change in the probability for an infinitesimal change in each independent variable (or the discrete change in the probability for dummy variable), evaluated at the mean of the independent variables.

As to the characteristics of policy decisions, the empirical results indicate that for decisions with large informational content (such as in the case of an interest rate change, as well as for large surprises), markets are less likely to change their direction during either the Q&A session or the introductory statement. In a similar fashion, strong market moves in response to the release or the introductory statement are also less likely to be corrected during the Q&A session, as can be seen from the negative coefficients estimated for the various residuals. These findings are revealing, as they suggest that there is less need for a fundamental clarification following communication that contains a lot of information.

Turning to the role of market and macroeconomic uncertainty, the results suggest that the market reactions during Q&A sessions

Table 5. Determinants of Directional Changes

	Intr. Statement vs. Release of Decision				Q&A vs. Release of Decision				Q&A vs. Intr. Statement			
	Actual		Control		Actual		Control		Actual		Control	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Proxies for News Content												
Characteristics of the Decision												
Absolute Surprise	−6.472**	2.811	—		−7.542***	2.108	—		−1.965	2.254	—	
Surprise Heterogeneity	1.096	2.455	—		1.927	2.630	—		−3.743	3.038	—	
Interest Rate Change	−0.029	0.246	—		−0.739***	0.087	—		−0.175	0.228	—	
Residuals												
Release of Decision	−60.864***	15.573	−32.610**	14.592	−49.898***	17.905	−24.074*	13.784	−9.620	11.619	22.814*	13.141
Introductory Statement	—		—		−18.969*	10.406	6.501	6.104	−29.766***	10.031	−5.076	7.201
Proxies for Uncertainty												
Prior Realized Volatility	−4.754	3.339	1.401	1.850	3.849	4.400	1.341	1.798	−2.672	3.793	−0.464	1.743
Macro Uncertainty												
Industrial Production	−0.049	0.166	0.019	0.082	0.453**	0.178	−0.118	0.083	0.613***	0.203	−0.138*	0.080
Inflation	0.668	1.454	0.359	0.624	2.871*	1.611	−0.165	0.644	0.741	1.430	−0.395	0.641
Money Growth	1.593***	0.466	−0.124	0.166	1.002***	0.400	0.023	0.161	0.100	0.374	0.082	0.161
# of Observations	53		188		53		188		53		188	
Pseudo R-Squared	0.380		0.026		0.484		0.025		0.270		0.029	

Notes: Results are based on estimation of model (4a)–(4c), using three-month Euribor futures, but with a discrete dependent variable, taking the value of one if the returns during two elements of the ECB's communication have a different sign, and zero otherwise. *, **, and *** denote significance at the 90 percent, 95 percent, and 99 percent level, respectively.

and introductory statements are more likely to lead to a market reversal in the presence of large macroeconomic uncertainty. The probability of market reversals is particularly elevated when comparing market movements in response to the decision and the Q&A session, highlighting that the Q&A session can serve as a useful tool for markets to clarify their opinions on the earlier decision. Importantly, no such pattern is found on control days, where macroeconomic uncertainty generally does not exert any effect on the probability of a reversal. The only exception suggests that, if anything, reversals are even less likely on non-press-conference days in the presence of macroeconomic uncertainty.

In sum, this section shows that holding a Q&A session gives the public the chance to ask clarifying questions, which appears to be especially relevant if there is large uncertainty about the macroeconomic environment in which monetary policy is operating.

6. Real-Time Effects of Press Conference Statements

The preceding sections have shown that the press conference contains valuable information for financial markets. But what is this additional information that is provided during press conferences, or more specifically, to what type of statements do financial markets react? To investigate these questions, the structure of the press conference is particularly helpful, as newswire services report in real time or market participants directly watch the broadcast of the press conference while at their trading desks. This allows us to trace the information flow and thus to investigate to what type of statements financial markets react predominantly.[18]

Table 7 in the appendix gives an impression about the way financial newswires report about the press conference. Because the snaps are recorded along with a time stamp, it is possible to identify the timing of the information flow. As mentioned in section 2, we distinguish the snaps according to their content, differentiating between

[18]This stands in contrast to the release of minutes or a press statement on the central bank's web site. As this is usually done through previous circulation to the press, albeit with an embargo time, financial newswires tend to prepare a number of snaps, which are then delivered simultaneously as soon as the embargo time has elapsed.

statements on the economic outlook, inflation, second-round effects, money growth, and interest rates. We create one time series for each of these categories; if a statement is classified accordingly, the time series for the corresponding category is allocated a "1" in the minute of the time stamp recorded by Reuters. For all other minutes, the variable is equal to zero. Our intention is to analyze the reaction of absolute returns, ticks per minute, and traded volumes to these variables. For that purpose, we will allow for at least one lag: if a Reuters snap is released toward the end of the minute, markets are most likely not reacting to this snap within this same minute. Hence, even under the assumption of near-instantaneous market responses, allowing for a lag is essential.

We include data from 14:30 to 15:45—i.e., the relevant time window for the press conference—and estimate the model for all Thursdays in the sample period, i.e., from September 2004 to July 2005. Finally, given the intraday patterns in the Euribor market (as seen in figures 3–5), it is essential to control for the time of day in such an analysis. Therefore, the regression model does include time dummies for each minute of this time window. The model is estimated as

$$y_t = \alpha + \sum_i (\beta_{1,i} x_{i,t} + \beta_{2,i} x_{i,t-1}) + \beta_{1,jobless} x_{jobless,t} + \beta_{2,jobless} x_{jobless,t-1} + \delta_t + \varepsilon_t, \quad (4)$$

where y_t denotes minute-by-minute absolute returns, number of ticks, or volume traded. $x_{i,t}$ denotes the variables for the different statement categories i as described above. $x_{jobless,t}$ stands for the absolute surprise component in the release of U.S. jobless claims at 14:30, measured by the difference between the released value and the median response in the Bloomberg survey. Finally, δ_t denotes a full set of time dummies, covering each minute from 14:30 to 15:45. The inclusion of a lag of the dependent variable does not alter the results in terms of significance of the estimated β-parameters. We thus decided against its inclusion, as the model without a lagged endogenous variable allows for an easier interpretation of the estimated parameters.

Table 6 reports the results, separately for absolute returns, ticks per minute, and volume traded in the three different panels. Three

Table 6. Market Reaction to Press Conference Statements

	Entire Press Conference		Introductory Statement		Q&A	
Absolute Return	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
U.S. Jobless Claims	0.000	0.003	0.001	0.003	—	—
Lagged Value	0.004	0.002	0.003	0.003	—	—
Economic Outlook	0.113***	0.025	0.121***	0.041	0.116***	0.034
Lagged Value	0.032	0.026	−0.018	0.042	0.085**	0.036
Inflation	0.068***	0.026	0.110***	0.037	0.036	0.041
Lagged Value	0.113***	0.026	0.143***	0.035	0.085**	0.040
Interest Rates	−0.026	0.046	—	—	−0.024	0.044
Lagged Value	0.231***	0.046	—	—	0.238***	0.045
Second-Round Effects	0.086*	0.046	0.067	0.059	—	—
Lagged Value	0.199***	0.051	0.230***	0.074	—	—
Money Growth	−0.031	0.040	−0.068	0.058	—	—
Lagged Value	0.091**	0.036	0.105**	0.047	—	—
R-square	0.049		0.098		0.030	
# of Observations	3,723		928		3,016	

(continued)

Table 6. (Continued)

	Entire Press Conference		Introductory Statement		Q&A	
Ticks per Minute	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
U.S. Jobless Claims	0.021	0.124	0.021	0.124	—	—
Lagged Value	0.211*	0.123	0.206*	0.124	—	—
Economic Outlook	2.434*	1.256	4.313**	1.728	1.289	1.642
Lagged Value	1.518	1.285	2.457	1.779	−0.138	1.702
Inflation	4.048***	1.303	3.437**	1.536	4.353**	1.953
Lagged Value	11.013***	1.269	6.442***	1.481	17.419***	1.912
Interest Rates	3.876*	2.264	—	—	3.970*	2.118
Lagged Value	12.690***	2.278	—	—	12.804***	2.145
Second-Round Effects	1.738	2.249	5.782**	2.461	—	—
Lagged Value	0.263	2.513	6.814**	3.111	—	—
Money Growth	4.704**	1.958	−2.393	2.450	—	—
Lagged Value	2.782***	1.774	10.549***	1.962	—	—
R-square	0.069		0.148		0.065	
# of Observations	3,723		928		3,016	

(continued)

Table 6. (Continued)

	Entire Press Conference		Introductory Statement		Q&A	
Volume	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
U.S. Jobless Claims	−2.357	10.737	−3.344	12.013	—	—
Lagged Value	7.576	10.642	0.319	12.012	—	—
Economic Outlook	108.511	109.018	−93.337	166.904	257.596*	138.142
Lagged Value	185.075*	111.532	441.008***	171.869	−87.970	143.132
Inflation	610.865***	113.019	841.212***	148.379	215.775	164.271
Lagged Value	789.608***	110.108	692.328***	143.083	1046.833***	160.841
Interest Rates	597.634***	196.436	—	—	610.282***	178.131
Lagged Value	1523.487***	197.618	—	—	1559.144***	180.417
Second-Round Effects	221.290	195.183	442.567*	237.714	—	—
Lagged Value	514.801**	218.013	861.098***	300.566	—	—
Money Growth	59.832	169.928	−216.735	236.686	—	—
Lagged Value	45.006	153.896	208.846	189.528	—	—
R-square	0.080		0.154		0.069	
# of Observations	3,723		928		3,016	

Notes: Results are based on estimation of model (4), testing for the effects of minute-by-minute newswire snaps on three-month Euribor futures. *, **, and *** denote significance at the 90 percent, 95 percent, and 99 percent level, respectively. Coefficients in panel 1 are multiplied by 100.

results are reported for each variable—once for the entire press conference, once for the introductory statement only, and once for the Q&A session only. For the last time window, the statements regarding second-round effects and money growth and their lags were discarded, as the data set contains fewer than ten entries for these.

The model comprises two types of controls—the surprise component in the U.S. jobless claim releases and the set of time dummies. A large number of time dummies are highly statistically significant, whereas no effect is found for the U.S. jobless claims. This might seem puzzling, especially given the spikes in trading at 14:30 on both press conference and non-press-conference days, which are clearly related to this data release. However, it is important to note that the release takes place at 14:30 each week, such that the time dummy for 14:30 and 14:31 will soak up any increase in market activity that is invariant across all days. The regressor $x_{jobless,t}$ contains the surprise component, which is estimated on top of the 14:30 and 14:31 effects. It is only this additional component that does not appear to affect the three-month Euribor futures in any significant fashion.

Looking at the response to the statement variables, there is clear evidence that returns, as well as trading activity, respond to the ECB's communication. The most robustly estimated effect, which is found across all three variables and for all three time windows, relates to statements about inflation—not surprisingly, given the importance of inflation data for the conduct of monetary policy. Adding up the contemporaneous and the lagged effect, a single statement about inflation affects returns by around 0.002 percent[19] (or changes implied future interest rates by around 0.2 basis points), leads to roughly fifteen additional trades, and increases the number of contracts traded by 1,400. Statements that relate directly to the discussion of policy rates in the Governing Council (which are never made during the introductory statement but are sometimes made in response to a question) have also clearly identified effects, on returns as well as on both measures of market activity. While the effects on returns and number of trades are about the same as those for inflation statements, substantially more trade volume is generated, with an increase of around 2,100 contracts (or €2 billion notional). Finally,

[19]Note that the parameters in panel 1 of table 6 are multiplied by 100, in order to enhance readability.

statements about possible second-round effects, money growth, and the economic outlook are found to be relatively influential, too, although the latter are particularly relevant if mentioned during the introductory statement and less so during the Q&A session.

7. Conclusions

Press conferences have recently become an important tool for several central banks to communicate monetary policy decisions to financial markets in real time. With several years of experience with press conferences among several central banks, it is now useful to evaluate this communication tool. This paper has exploited the experience of the ECB with press conferences, analyzing in particular (i) to what extent they provide systematic information in addition to the release of policy decisions and (ii) specifically, whether the press conferences fulfill a clarification role for financial markets.

The results of the paper indicate that press conferences add substantial information to the release of the decisions themselves, often exerting an even larger effect on financial markets than the release of the decisions. One of the central findings of the paper is that the information content of the press conference appears to be closely related to the characteristics of a given decision, as the introductory statement adds information in a systematic manner when the policy decision has been relatively unexpected. In particular, press conferences are found to be especially useful when there is a high degree of macroeconomic uncertainty. Under such circumstances, market participants are more likely to seek guidance from central bank communication and show a more muted reaction to the release of the decisions but a larger response to press conferences, and in particular Q&A sessions, as these provide clarification.

The paper has also analyzed what type of information is particularly relevant for financial markets, using data of minute-by-minute newswire snaps. It is specifically statements about inflation as well as statements related directly to the discussion of policy rates in the Governing Council that exert the largest and most systematic impact on financial markets during the press conference. Statements about second-round effects, the economic outlook, and money growth also influence financial markets, though their effects are less significant statistically.

In sum, the paper suggests that press conferences can provide a useful tool in explaining monetary policy to the public, in particular because the sequence of events allows for a useful flow of information, whereby later parts of the communication sequence can react to the information content of earlier parts, a pattern most evident in the form of the clarification that Q&A sessions can provide. Given the importance of a common understanding between the public and the central bank for the effectiveness of monetary policy, this advantage cannot be overemphasized. However, the focus on the ECB's case in this paper leaves open the question of how other communication tools perform in comparison. We leave this important policy question for future research.

Appendix

Table 7. Reuters Snaps during the ECB's Press Conference in November 2004

Date	Time	Snap
04/11/04	14:32	TRICHET—CPI A WORRISOME DEVELOPMENT
04/11/04	14:32	TRICHET—NO EVIDENCE INFLATION PRESSURES PICKING UP IN EURO ZONE
04/11/04	14:32	TRICHET—UPSIDE RISKS TO PRICE STABILITY, NEED STRONG VIGILANCE
04/11/04	14:33	TRICHET—SHORT-TERM ECON INDICATORS MORE MIXED
04/11/04	14:33	TRICHET—INDICATORS STILL POINT TO GROWTH IN 2005
04/11/04	14:33	ECB'S TRICHET SAYS EURO AREA EXPORTS SHOULD STILL BENEFIT FROM GLOBAL DEMAND NEXT YEAR
04/11/04	14:34	TRICHET—STILL SCOPE FOR STRONGER PRIVATE CONSUMPTION
04/11/04	14:34	TRICHET—OUTLOOK SURROUNDED BY UNCERTAINTY, MAINLY FROM OIL
04/11/04	14:34	TRICHET—MAGNITUDE, NATURE OF OIL SHOCK DIFFERENT FROM PAST
04/11/04	14:35	TRICHET—RECENT OIL PRICE RISES STILL SIZABLE ADVERSE SHOCK TO EURO ZONE
04/11/04	14:35	TRICHET—MORE OIL RISES COULD DAMPEN GROWTH
04/11/04	14:36	ECB'S TRICHET SAYS OCTOBER CPI SAW STRONG JUMP, OIL PRICE HAD DIRECT IMPACT
04/11/04	14:36	TRICHET—OIL PRICES MAY FEED THROUGH ECON, GENERATE INDIRECT EFFECTS
04/11/04	14:36	TRICHET—DATA SO FAR DO NOT SUGGEST STRONGER UNDERLYING CPI PRESSURES
04/11/04	14:36	TRICHET—WAGE MODERATION EXPECTED TO CONTINUE, GIVEN MODERATE GROWTH

(continued)

Table 7. (Continued)

Date	Time	Snap
04/11/04	14:37	TRICHET—RISKS LINKED TO OIL PRICES, TAXES, POTENTIAL SECOND-ROUND EFFECTS
04/11/04	14:37	TRICHET—DOWNTREND IN M3 APPEARS TO HAVE HALTED IN RECENT MONTHS
04/11/04	14:37	ECB'S TRICHET SAYS M3 DEVELOPMENTS REFLECT LOW INTEREST RATES
04/11/04	14:38	TRICHET—LOW INT RATES FUELLING PRIVATE CREDIT DEMAND
04/11/04	14:38	TRICHET—LOAN DEMAND NOW MORE BROADLY BASED
04/11/04	14:38	TRICHET—MORE CASH IN EURO ZONE THAN NEEDED TO FINANCE CPI-FREE GROWTH
04/11/04	14:39	TRICHET—EXCESS CASH COULD POINT TO INFLATION AHEAD, BOOST ASSET PRICES
04/11/04	14:39	TRICHET—UNDERLYING INFLATION CONTAINED BUT MEDIUM-TERM RISKS NEED TO BE MONITORED
04/11/04	14:39	ECB'S TRICHET SAYS NEED STRONG VIGILANCE ON RISKS TO PRICE STABILITY
04/11/04	14:40	TRICHET—SEE SOME ENCOURAGING SIGNS THAT GOVTS PLANNING TO CORRECT BUDGET DEFICITS
04/11/04	14:40	TRICHET—BUT FISCAL IMBALANCES ELSEWHERE COULD BE ON THE RISE
04/11/04	14:41	TRICHET—COMPLYING WITH BUDGET TARGETS WILL BUILD CONFIDENCE, HELP UPSWING
04/11/04	14:41	TRICHET—ECB WARNS AGAINST CHANGES TO STABILITY PACT, EXCESS DEFICIT PROCEDURE
04/11/04	14:42	TRICHET—FISCAL CONSOLIDATION PLANS SHOULD BE PART OF STRUCTURAL REFORM EFFORT
04/11/04	14:42	ECB'S TRICHET SAYS STRUCTURAL REFORM CRUCIAL TO RAISE GROWTH, EMPLOYMENT
04/11/04	14:44	TRICHET—ASKED IF ECB HAS BIAS, SAYS UPWARD RISKS TO PRICE STABILITY "AUGMENTING"
04/11/04	14:44	TRICHET—"WE WON'T LET SECONDARY EFFECTS MATERIALISE"
04/11/04	14:45	TRICHET—REAFFIRMS G7 STATEMENT ON CURRENCIES
04/11/04	14:45	TRICHET—G7 STATEMENT CAPTURES "PRESENT SENTIMENT"
04/11/04	14:46	TRICHET—DISORDERLY MOVEMENTS IN FX RATES UNDESIRABLE FOR GROWTH
04/11/04	14:46	ECB'S TRICHET SAYS U.S. POLICY REMAINS ONE OF "STRONG DOLLAR"
04/11/04	14:49	TRICHET—"CLEAR" THAT PRESENT PICKUP IN CPI WILL CONTINUE FOR SOME MONTHS, MAY INCREASE
04/11/04	14:50	TRICHET—WILL RETURN TO PRICE STABILITY "AT A CERTAIN MOMENT"
04/11/04	14:50	TRICHET—DECLINES TO FORECAST AVERAGE CPI FOR 2005
04/11/04	14:51	TRICHET—OIL IMPACT ON INFLATION UPWARD, ON GROWTH DOWNWARD
04/11/04	14:53	TRICHET—ASKED ON EURO FX RATE, REPEATS EXCESS VOLATILITY, DISORDERLY MOVES, UNDESIRABLE

(continued)

Table 7. (Continued)

Date	Time	Snap
04/11/04	14:54	ECB'S TRICHET SAYS SEES ASSET INFLATION IN REAL ESTATE IN A NUMBER OF ECONOMIES
04/11/04	14:56	TRICHET—ON U.S. TWIN DEFICITS, ECB SAYS U.S. NEEDS TO CORRECT SAVINGS PROBLEM
04/11/04	14:56	TRICHET—EUROPE NEEDS TO TACKLE STRUCTURAL REFORMS AS PART OF REDUCING IMBALANCES
04/11/04	14:56	TRICHET—NEED PROGRESSIVE, EFFICIENT CORRECTION OF LACK OF SAVINGS IN U.S.
04/11/04	14:58	TRICHET—APPROPRIATE FOR CHINA TO USE "MARKET ECONOMY WEAPONS" TO COOL ECONOMY
04/11/04	14:59	TRICHET—NUMBER OF ASIA CURRENCIES COULD APPRECIATE IN PROGRESSIVE, ORDERLY MANNER
04/11/04	15:03	ECB'S TRICHET—ECB READY TO PREVENT SECOND-ROUND EFFECTS
04/11/04	15:10	TRICHET—IN PERMANENT CONTACT WITH U.S., BUT NONE SINCE U.S. ELECTION
04/11/04	15:11	TRICHET—WHAT IS GOOD FOR CHINA ECONOMY IS GOOD FOR GLOBAL ECONOMY
04/11/04	15:11	TRICHET—DOES NOT SEE RISK OF STAGFLATION FOR EURO ECONOMY NOW
04/11/04	15:12	TRICHET—"WE HAVE IN HAND" THE DELIVERANCE OF PRICE STABILITY OVER TIME
04/11/04	15:13	TRICHET—WILL CONTINUE TO SEE GROWTH CLOSE TO POTENTIAL IN EURO AREA
04/11/04	15:17	ECB'S TRICHET SAYS DOES NOT LIKE "ONE OFF" FISCAL MEASURES AS A RULE

References

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys. 2003. "Modeling and Forecasting Realized Volatility." *Econometrica* 71 (2): 579–626.

Andersson, M. 2008. "Using Intraday Data to Gauge Financial Market Responses to Fed and ECB Monetary Policy Decisions." Forthcoming, *International Journal of Central Banking.*

Andersson, M., H. Dillen, and P. Sellin. 2006. "Monetary Policy Signaling and Movements in the Term Structure of Interest Rates." *Journal of Monetary Economics* 53 (8): 1815–55.

Andersson, M., L. J. Overby, and S. Sebestyén. 2009. "Which News Moves the Euro Area Bond Market?" *German Economic Review* 10 (1): 1–31.

Berger, H., J. de Haan, and J.-E. Sturm. 2006. "Does Money Matter in the ECB Strategy? New Evidence Based on ECB Communication." CESifo Working Paper No. 1652.

Berger, H., M. Ehrmann, and M. Fratzscher. 2006. "Monetary Policy in the Media." ECB Working Paper No. 679 (September).

Bernanke, B. 2004. "Fedspeak." Remarks at the meetings of the American Economic Association, San Diego, California, January 3. Available at www.federalreserve.gov/boarddocs/speeches/2004/200401032/default.htm.

Bernanke, B., and K. Kuttner. 2005. "What Explains the Stock Market's Reaction to Federal Reserve Policy?" *Journal of Finance* 60 (3): 1221–57.

Blinder, A. S. 1998. *Central Banking in Theory and Practice*. Cambridge, MA: MIT Press.

Bomfim, A. N. 2003. "Pre-announcement Effects, News Effects, and Volatility: Monetary Policy and the Stock Market." *Journal of Banking and Finance* 27 (1): 133–51.

Brand, C., D. Buncic, and J. Turunen. 2006. "The Impact of ECB Monetary Policy Decisions and Communication on the Yield Curve." ECB Working Paper No. 657.

Cochrane, J. H., and M. Piazzesi. 2002. "The Fed and Interest Rates: A High-Frequency Identification." *American Economic Review Papers and Proceedings* 92 (2): 90–95.

de Haan, J., F. Amtenbrink, and S. Waller. 2004. "The Transparency and Credibility of the European Central Bank." *Journal of Common Market Studies* 42 (4): 775–94.

Ehrmann, M., and M. Fratzscher. 2003. "Monetary Policy Announcements and Money Markets: A Transatlantic Perspective." *International Finance* 6 (3): 309–28.

———. 2004. "Taking Stock: Monetary Policy Transmission to Equity Markets." *Journal of Money, Credit, and Banking* 36 (4): 719–38.

———. 2007. "Communication by Central Bank Committee Members: Different Strategies, Same Effectiveness?" *Journal of Money, Credit, and Banking* 39 (2–3): 509–41.

Evans, M., and R. Lyons. 2005. "Do Currency Markets Absorb News Quickly?" *Journal of International Money and Finance* 24 (2): 197–217.

Faust, J., J. H. Rogers, S-Y. Wang, and J. H. Wright. 2007. "The High-Frequency Response of Exchange Rates and Interest Rates to Macroeconomic Announcements." *Journal of Monetary Economics* 54 (4): 1051–68.

Fleming, M., and E. Remolona. 1999. "Term Structure of Announcement Effects." Federal Reserve Bank of New York Staff Report No. 76.

Gaspar, V., G. Perez-Quiros, and J. Sicilia. 2001. "The ECB Monetary Policy Strategy and the Money Market." *International Journal of Finance and Economics* 6 (4): 325–42.

Geraats, Petra. 2002. "Central Bank Transparency." *The Economic Journal* 112 (483): 532–65.

Goodhart, C. 2005. "Dear Jean-Claude." *Central Banking* 16 (1): 32–36.

Gropp, R., and A. Kadareja. 2006. "Stale Information, Shocks and Volatility." ECB Working Paper No. 686.

Gürkaynak, R., B. Sack, and E. Swanson. 2005. "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements." *International Journal of Central Banking* 1 (1): 55–94.

Guthrie, G., and J. Wright. 2000. "Open Mouth Operations." *Journal of Monetary Economics* 46 (2): 489–516.

Hartmann, P., M. Manna, and A. Manzanares. 2001. "The Microstructure of the Euro Money Market." *Journal of International Money and Finance* 20 (6): 895–948.

Heinemann, F., and K. Ullrich. 2007. "Does It Pay to Watch Central Bankers' Lips? The Information Content of ECB Wording." *Swiss Journal of Economics and Statistics* 143 (2): 155–85.

Issing, Otmar. 2005. "Communication, Transparency, Accountability: Monetary Policy in the Twenty-First Century." *Review* (Federal Reserve Bank of St. Louis) 87 (2, Part 1): 65–83.

Kohn, D. L., and B. P. Sack. 2004. "Central Bank Talk: Does It Matter and Why?" In *Macroeconomics, Monetary Policy, and Financial Stability: A Festschrift in Honour of Charles Freedman*, 175–206. (Proceedings of a conference held by the Bank of Canada, June 2003.) Ottawa: Bank of Canada.

Kuttner, K. 2001. "Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market." *Journal of Monetary Economics* 47 (3): 523–44.

Reeves, R., and M. Sawicki. 2007. "Do Financial Markets React to Bank of England Communication?" *European Journal of Political Economy* 23 (1): 207–27.

Reinhart, V., and B. Sack. 2006. "Grading the Federal Open Market Committee's Communications." Mimeo, Board of Governors of the Federal Reserve System (January).

Rigobon, R., and B. Sack. 2004. "The Impact of Monetary Policy on Asset Prices." *Journal of Monetary Economics* 51 (8): 1553–75.

Rosa, C., and G. Verga. 2007. "On the Consistency and Effectiveness of Central Bank Communication: Evidence from the ECB." *European Journal of Political Economy* 23 (1): 146–75.

Rosenberg, I. 2007. "Changes in Monetary Policy Communication." Speech given at a meeting of the Stockholm-Djurgården Rotary Club, Stockholm, Sweden, May 11. Available at http://www.riksbank.com/templates/Page.aspx?id=24741.

Thornton, D. L. 1998. "Tests of the Market's Reaction to Federal Funds Rate Target Changes." *Review* (Federal Reserve Bank of St. Louis) 80 (Nov/Dec): 25–36.

Woodford, M. 2005. "Central-Bank Communication and Policy Effectiveness." Paper presented at the Federal Reserve Bank of Kansas City Symposium on The Greenspan Era: Lessons for the Future, Jackson Hole, Wyoming, August 25–27.

The Interest Rate Conditioning Assumption*

Charles Goodhart
Financial Markets Group, London School of Economics

A central bank's forecast must contain some assumption about the future path for its own policy-determined short-term interest rate. I discuss the advantages and disadvantages of the three main alternatives:

(i) constant from the latest level
(ii) as implicitly predicted from the yield curve
(iii) chosen by the monetary policy committee (MPC)

Most countries initially chose alternative (i). With many central banks having planned to raise interest rates at a measured pace in the years 2004–06, there was a shift to (ii). However, Norway, and now Sweden, has followed New Zealand in adopting (iii), and the United Kingdom has also considered this move. So this is a lively issue.

JEL Codes: E47, E52, E58.

1. Introduction

A central bank's forecast must contain some assumption about the likely future path for its own policy-determined short-term interest rate. Most of those central banks that have publicly reported their procedures in this respect have in the past assumed that interest rates would remain unchanged from their present level, e.g., in Sweden, until recently,[1] and in the United States (at least most of

*My thanks are due to Peter Andrews, David Archer, Oriol Aspachs, Charlie Bean, Jarle Bergo, Hyun Shin, Lars Svensson, Bent Vale, Mike Woodford, my two referees, and the members of the Bank of England seminar on August 3, 2005, for helpful comments. The views, and remaining errors, in this paper remain, however, my own responsibility.

[1]It was reported, e.g., in the *Financial Times* Lex column, January 30, 2007, in the article entitled "Central Bank Forecasting," that Sweden had joined the group (plus New Zealand and Norway) giving conditional forecasts of the expected future path of their own policy-determined interest rates.

the time) (for Sweden, see Berg, Jansson, and Vredin 2004, and Jansson and Vredin 2003; for the United States, see Boivin 2004, Reifschneider, Stockton, and Wilcox 1997, and Romer and Romer 2004). The United Kingdom was amongst this group from the Bank of England's first Inflation Report, at the end of 1992, until May 2004; then in August 2004 it shifted to the use of the forward short rates that are implied by the money-market yield curve.[2] But Deputy Governor Lomax stated (2007) that the Bank of England was considering joining the small group of countries (New Zealand, Norway, and Sweden) that are explicitly reporting their own expectations for the future path of interest rates. So, in this paper the focus will be on the question of how a monetary policy committee (MPC) does, and should, choose (condition) a future time path for its own policy variable, the officially determined short-term interest rate.

There are two main purposes for such forecasting exercises: the first is as an aid to the policy decision itself, which is to choose the current level of official short-term interest rates; the second is to communicate to the general public both an explanation of why the official rate was changed and an indication of how the MPC views future economic developments. The manner in which these two purposes may be linked depends in some large part on the institutional detail of the manner in which each individual MPC has been established.

For example, prior to its being given operational independence in May 1997, the Bank of England's inflation forecast in its Inflation Report (starting in 1993) was intended to be an aid to the choice of interest rates taken by the Chancellor of the Exchequer (see Goodhart 2001b). Since the decision remained with the Chancellor, however, the Bank felt that it should not be seen to be pushing the Chancellor to follow any particular path for interest rates. So its forecast was conditioned on a neutral assumption, that interest rates remained constant (in nominal terms) from whatever level they had previously reached.

[2]In fact, it used *both* conditioning assumptions for many years before 2004, but the constant interest rate assumption was given clear precedence. Since August 2004, it has continued to use both conditioning assumptions, but now the money-market rate curve is given the greater emphasis (see Lomax 2005).

In order to provide a basis for such inflation forecast(s), which then forms one of the main inputs into the current interest rate decision, the only strong requirement is that the conditioning assumption for the future path of short-term policy rates is not too patently out of line with what the decision makers, and the markets, believe will actually happen. For simplicity, most MPCs initially chose constant future policy interest rates, from the latest available level, as their main framing assumption. Occasionally, such an assumption would have been grossly at odds with perceived reality, as in the case of the United States from 2004 until early 2006, when the explicit position of the Federal Open Market Committee (FOMC) was for there to be a "measured increase" in policy rates over time. In that case, the Greenbook conditioning assumption, which has also been *usually* for constant rates,[3] is widely believed to have been changed, but the degree of secrecy, and length of lag before publication (five years), means that we will not have confirmation of this for some time.

Of course, in addition to the basic conditioning assumption, MPC members can ask for alternative scenarios to be run, involving differing conditional time paths. There can be as many such simulations run as the resources, time, and technical skills of the Bank staff allow. But, for the purposes of communication, only *one* forecast is generally published, albeit now often including probability distributions (fan charts). On all this, see Edey and Stone (2004).

A crucial distinction, however, lies between those MPCs that just publish a "staff forecast" giving the forecast conditioned on the staff's own (standard) interest rate assumption, and those where the forecast is issued under the aegis of the MPC, or a decision-making Governor. Examples of the former are the European Central Bank (ECB) and the FOMC; examples of the latter are the United Kingdom's MPC, Norway, Sweden, and New Zealand.

Requirements for the former are less restrictive than for the latter. Thus, MPCs presenting a staff forecast need not even update that forecast to incorporate the actual subsequent decision. The publication of a staff forecast, on a standard conditioning assumption, then simply reveals a key input into the decision-making procedure.

[3]This was not always so. It was upward sloping in 1994.

It is, in a sense, a simulation, not a true forecast, and should be interpreted as such.

The situation is different when what is to be presented is a forecast for which the MPC (Governor) actually takes responsibility. This crucial change in context was not, perhaps, fully appreciated when the Bank of England was given operational independence, and the UK MPC was formed, in May 1997. Then the constant interest rate assumption, which had been appropriate in the earlier regime, was simply continued, without much consideration or public discussion.

2. Arguments against a Constant Interest Rate Assumption

The strongest single argument against the assumption of a constant future nominal short-term interest rate path in a proper forecast, as contrasted with a "staff forecast," or simulation, is that this is often *not* what the central bank itself nor the money market expect to happen. The money-market yield curve is only occasionally approximately flat out to the forecast horizon (which for the purpose of this exercise we take to be eight quarters ahead).[4] Perhaps even more important, there have been periods when a central bank has been clearly signaling that it expected future changes in its policy-determined interest rates. The expectation of a "measured" rate of increase in U.S. interest rates in 2004–05 is a case in point. But such signaling was also apparent in the United Kingdom in early 2004. It is, to say the least, inconsistent to have the central bank give one message in words and then base its published forecast on quite a different assumption.

Even when it is just a staff forecast, or simulation, rather than an MPC forecast, too glaring a deviation between conditioning assumption and actual expectations reduces the role of such a simulation, either as an input into policy decisions or as a means of communication with the public. If the staff forecast should be based on a

[4]In August 2004 the MPC in the United Kingdom extended the horizon recorded in the forecasts (for inflation and output growth) to three years, but the surrounding text tended to indicate that the two-year horizon remained the chief focus of attention. Again, see Lomax (2005).

conditioning assumption for the future path of policy rates significantly different from that expected by the decision makers, it will be harder for the latter to reach a sensible, informed view for the current decision on policy rates. It would then also be somewhat more difficult to explain that latter decision to the public in terms of expected future inflation (and output gaps), even if the staff forecast is not published. The difficulty would become much more acute if the staff forecast was then to be published. With MPC forecasts being published, any serious deviation between the actual expectations of the MPC and the conditioning assumptions for the future path of policy rates could lead to major problems in communicating with the public.

In particular, when the policy interest rate is cyclically high—or low, as it patently was in many countries after 2001—extrapolating the current level of interest rates into the future will give implausible results and cannot therefore be either a sensible basis for internal decisions or a fruitful means of communication with the private sector. Adolfson et al. (2005, 1) used a DSGE model to simulate monetary policy in the euro area and found that "in the latter part of the sample (1998:Q4–2002:Q4) ... the constant interest rate assumption has arguably led to conditional forecasts at the two-year horizon that cannot be considered economically meaningful during this period."

3. Should an MPC Forecast the Future Time Path of Its Own Official Rate?

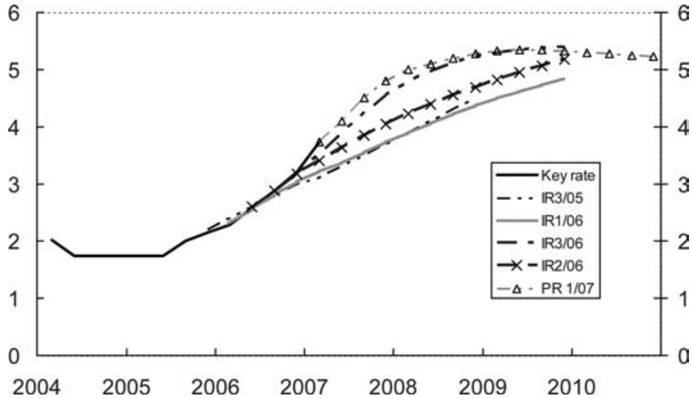
The main alternative in the academic literature, which several economists have been advocating (e.g., Svensson 2003, 2004 and Woodford 2004), is to base the conditioning assumption on a specific nonconstant *forecast* made by the Bank or by its MPC. But this also has its drawbacks. While an MPC might be quite willing to agree and to endorse a *general* direction of likely future change (as in the FOMC "bias" reports or the ECB's standard vocabulary), it would generally be much less happy to commit itself to a specific, quantitative path, although this is what *has* been done in New Zealand, and its relatively untroubled acceptance there influenced Svensson, who wrote a report on their procedures (Svensson 2001). This has also been done since 2006 in Norway, and since 2007 in Sweden. Lomax (2007) reported that the UK MPC was also considering this step.

In New Zealand the responsibility for hitting the inflation target rests on the Governor of the Reserve Bank personally. So he (as yet there have been no female Governors there) can also decide upon the form and nature of the published forecast, including the conditioning assumptions. It is difficult enough for an MPC to agree on the selection of the policy rate to hold until the next meeting, when the range of feasible and sensible options is quite limited (and that range has been greatly reduced by the implicit, but now general, convention that interest rate changes should always be in multiples of 25 basis points); it would be a quantum leap more difficult to get such a committee to agree on a single path for the next n quarters, when the potential range of feasible/sensible options widens dramatically (see Mishkin 2004). The procedure for adopting a specific forecast future path for interest rates is made easier when a Governor has sole responsibility (New Zealand) or the relevant committee is small, as in Norway (where the Governor usually has a decisive role) and Sweden.

Assuming that an MPC could agree, or find a procedure for agreeing, on such a forecast for the time path of future interest rates (Svensson has suggested taking the median of individually decided preferred paths), this would almost certainly have to be published. In view of the current ethos of transparency, it would hardly be acceptable to state that the forecast was based on a nonzero conditioning assumption, but that the public is not to be told what this was (though on some occasions the Federal Reserve staff have based their Greenbook forecasts on a nonconstant rate assumption without any clear indication of what that assumption was being available to the public, since such forecasts are protected from public inspection by the five-year lag in publication).

If an MPC's nonconstant forecast was to be published, there is a widespread view, in most central banks, that it would be taken by the public as more of a commitment and less of a rather uncertain forecast than should be the case. That concern can, however, be mitigated by producing a fan chart of possible interest rate paths, rather than a point estimate, and/or by publishing additional scenario paths. No doubt, though, measuring rulers and magnifying glasses would be used by private-sector observers to extract the central tendency. Examples of recent published forecasts for Norway and New Zealand are given in figures 1 and 2. Once there was a

Figure 1. Key Policy Rate and Projections of the Key Rate Since Autumn 2005 in Norway



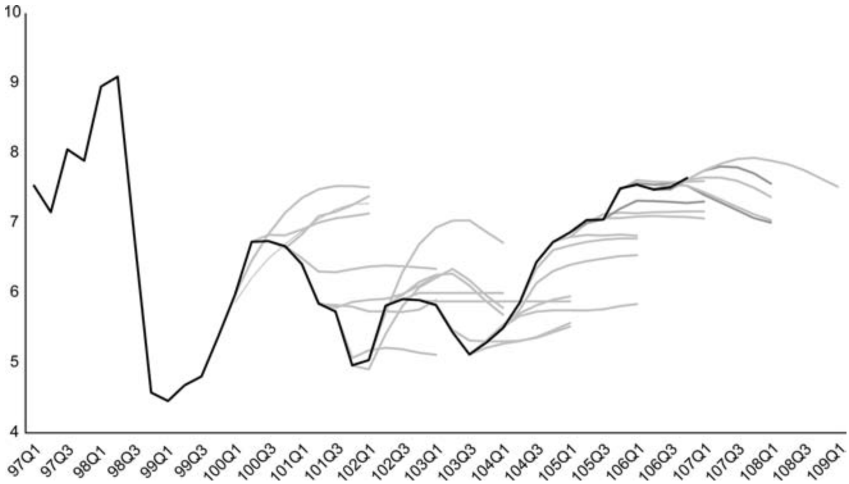
Source: Norges Bank.

published central tendency, then this might easily influence the private sector's own forecasts more than its own inherent uncertainty warranted, along lines analyzed by Morris and Shin (1998, 2002, 2004).[5] Likewise, when new, and unpredicted, events occurred and made the MPC want to adjust the prior forecast path for interest rates, this might give rise to criticisms, ranging from claims that the MPC had made forecasting errors to accusations that they had reneged on a (partial) commitment.

Lars Svensson and some other academics respond that this worry implies that MPCs regard participants in financial markets as unsophisticated and incapable of understanding the concept of a conditioning assumption. Moreover, there have been few, if any, recorded problems in New Zealand; some recent Norwegian concerns are discussed later on here. Moreover, it could be argued that having to explain the reasons why it has deviated from its prior forecast could be a good discipline for the central bank. But these countries have small financial systems, clearly dependent on international developments; reactions there may differ from those in larger countries. Be

[5]There has been a continuing debate between Svensson and Morris, Shin, and Tong on the necessary conditions under which transparency may, or may not, be damaging to social welfare. See Svensson (2005) and Morris, Shin, and Tong (2005).

Figure 2. Key Policy Rate and Projections of the Key Rate Since March 2000 in New Zealand



Source: David Archer.

that as it may, most members of MPCs have been reluctant to move to a specific forecast for a future time path for interest rates.

One of my (anonymous) referees added that the appropriate path of the policy rate can also depend, in part, on a wide range of other financial variables (equity prices, risk spreads, currently the likelihood and effect of a "credit crunch," and so on) or, depending on the sophistication of the model used, risk premiums on the various assets (equities, corporate bonds, and so on). Thus, to allow the public to make sense of the projected policy path, the central bank might, at least at times, have to provide information on these other variables. So, for example, in the late 1990s, some of the (publicly released) Greenbooks noted that the projected path for policy was fairly flat because of an assumed leveling out in stock prices. Is that really something that the central bank would like to say publicly? Moreover, such financial variables could easily turn out differently than anticipated (e.g., the 1987 NYSE crash or the 2007 credit-market freeze), but the central bank would likely intend in such circumstances to offset the effects on the real economy by adjusting policy. So, in a sense, the policy assumption is more tentative and more subject to change than the projections for output and inflation.

A related, but reverse, argument is that it would not be the private sector, but *the MPC itself* that might place too much weight on an explicit forecast path. Thus, having given a forward projection, an MPC might feel pressured to stick to it, even when circumstances had changed. This was the gist of an editorial in the *Financial Times* (December 7, 2006, p. 20) entitled "Giving a Wrong Signal."[6] This editorial included the following passage:

> However, the market is far more interested in detecting any hints that Jean-Claude Trichet, the ECB president, might give regarding monetary policy in 2007. Mr Trichet's communication strategy has reached a level of comical transparency: a mention of "vigilance" signals a rise in the following month, while "monitoring closely" means it will happen two or three months hence.
>
> Such signposting does have some merits. But pre-announcing interest rate decisions also entails an obvious loss of flexibility. And in the increasingly uncertain global outlook of 2007 this flexibility will be needed. . . . The economic outlook is uncertain. Mr Trichet should make sure his language reflects this.

4. Using an Implied Market-Based Forecast for Future Official Rates

Caught between the lack of credibility (at least on some occasions) of a constant rate assumption and the problems of adopting an MPC chosen time path for interest rates, the move by the UK MPC to adopt the estimated future path *as estimated by the market* for its

[6]Ehrmann and Fratzscher (2007) report that the Federal Reserve's policy directives before 1999, when they were unpublished and for internal use only, were a much less accurate predictor of subsequent policy moves than after May 1999, when they "were targeted at an external audience" (see especially footnote 7, p. 189). While there may be several other reasons for this, such behavior is consistent with the possibility that publication of future plans acts as a commitment device for carrying them out later. Exactly how far it is desirable for an MPC to commit itself to a future path for interest rates, in a world of uncertainty, remains uncertain. For arguments in favor of some such commitment, see Woodford (2003, ch. 7); for arguments against, see Issing (2005), as quoted by Ehrmann and Fratzscher (2007, 222–23).

conditioning assumption could be seen as a brilliant compromise that got around the worst features of both the other two alternatives. Given the normal assumptions of rational expectations and efficient markets, the market's forecast ought to be credible, yet its adoption in the forecasting procedure required no decision procedure in the MPC itself and committed them to *nothing*, a master stroke indeed. The change in procedure did not at the time cause much discussion or elicit any criticism (that I saw). There may, however, be some drawbacks to this new approach, which need to be considered. One issue is the dynamic implications of adopting a market forecast; a second is how far the market forecast has had a good track record. The latter remains the subject of my further, ongoing research, which Wen Bin Lim and I intend to undertake.

Yet another of the criticisms raised against the *constant* interest rate forecast is that, if maintained too long, it would lead to Wicksellian instability. Indeed in medium-run simulations at the Bank of England extending much beyond the prior two-year horizon, the constant two-year rate assumption had to be linked into a Taylor-type reaction function to prevent nonsensical trends developing as the horizon passed beyond two years. But, up to the two-year horizon, there did not seem to be any practical, empirical problem with this assumption, as also noted in Edey and Stone (2004).

On the other hand, the assumption of constant forward policy-determined interest rates imposed a strong discipline on the MPC that may be considered to be strongly beneficial (see Goodhart 2001a). Because of the UK MPC's inbuilt dislike of reporting inflation failing to come back close to target at their focus horizon of seven or eight quarters hence, this assumption virtually forced the MPC to take immediate, and sufficient, action to counter and remove any perceived threat to inflation stability as soon as it appeared. This behavioral trait was documented in several recent papers (Goodhart 2004, 2005). In my view, the main cause of endemic inflation in earlier decades had been the syndrome of "too little, too late" in a context of great uncertainty, a trait which could be viewed as a version of time inconsistency. So any procedure that, more or less, forced the decision makers into prompt corrective action was to be supported and encouraged.

What will be the dynamic implications for the new market-based forecasting mechanism? It is, to say the least, an incestuous exercise. The market is trying to guess what the authorities will do, and their guess is then incorporated as the conditioning assumption to the initial forecast on which, in part, the MPC bases its decision.

Clearly there are no problems when the MPC's current decision has been (largely) predicted by the market and the resultant forecast shows inflation reverting satisfactorily to target. But what if the MPC's forecast should indicate (given the current decision *and* the implied money-market yield curve) that inflation would still be tending to overshoot (undershoot) the target, especially, but not only, at the key horizon?[7] Then (as emphasized by Bank of England economists) the *publication* of that deviation would influence *expectations* of market participants in the desired direction and lead to an appropriate rise (fall) in *future expected* rates and hence in longer-term interest rates. Then, movements in longer-term interest rates will affect the economy more widely. Thus, goes the argument, the Bank now has effectively two instruments—its current interest rate decision *and* its separate ability to influence expected future interest rates.[8] The latter is not, however, an instrument that the

[7] Owing to lags in the transmission mechanism whereby interest rates affect the economy, any attempt to vary such rates to bring inflation back to target quickly would lead to (instrument) instability. Instead, the authorities tend to focus on a crucial longer horizon for restoring inflation to target. In the United Kingdom, that key horizon has been about seven or eight quarters from the forecast date.

[8] This is closely similar to the analysis in Gürkaynak, Sack, and Swanson (2005, 86–87), in which they state the following:

> Do central bank actions speak louder than words? We find that the answer to this question is a qualified "no." In particular, we find that viewing the effects of FOMC announcements on financial markets as driven by a single factor—changes in the federal funds rate target—is inadequate. Instead, we find that a second policy factor—one not associated with the current federal funds rate decision of the FOMC but instead with statements that it releases—accounted for more than three-fourths of the explainable variation in the movements of five- and ten-year Treasury yields around FOMC meetings.
>
> We emphasize that our findings do not imply that FOMC statements represent an *independent* policy tool. In particular, FOMC statements likely exert their effects on financial markets through their influence on financial market expectations of *future* policy actions. Viewed in this light, our results do not indicate that policy actions are secondary so much as that their influence comes earlier—when investors build in expectations of those actions in response to FOMC statements (and perhaps other events, such as speeches and testimony by FOMC members).

Bank can vary at will. If the Bank's forecast was ever suspected of being manipulated to achieve a market effect, it would lose all credibility. The Bank is forced to give its best, most truthful, forecast. Indeed, moving from a "one-instrument regime" (only operating on short-term interest rates) to a "two-instrument regime" (operating on both short-term interest rates and future interest rate expectations) might allow the central bank to vary the short-term rate less than otherwise. This is a point that has been emphasized by Woodford (2003 and 2005, for example).

That is an argument that I accept, *up to a point.* If the resulting deviation of inflation from target, as shown in the Inflation Report, is large, especially at the key horizon of seven or eight quarters hence, and/or continuously worsening, it would raise public queries as to why no action had already been taken to deal with the perceived inflationary (deflationary) threat. While it may be possible to give answers to this, the extent to which the MPC has been prepared to allow forecast inflation to deviate from target, especially at the crucial horizon of around seven or eight quarters, has been historically small.

However, this is not an argument that the Norges Bank has found acceptable. They state that the main reason for switching to a specific forecast path in 2006 was that the path of future rates implied by the market yield curve was then too flat and low to be consistent with a return to normal conditions.[9] The Bank believed that future policy rates would, and should, be rising. Rather than publish a forecast based on market rates implying an increasing boom and incipient inflationary pressures, based on a market rate forecast, they preferred to publish a forecast of their own conditional expectations. This was an important factor in their decision to base their forecast and published Inflation Report on their own future expected path for policy rates.

5. Market Reactions to Surprises in the Forecast

Moreover, with a market-based forecast, what happens if the MPC's current decision surprises the market, in the sense that it has not (or

[9]This information is from a personal discussion on January 25, 2007.

has only partly) been previously expected? Clearly an unexpected change in direction will have greater impact than an unexpected change in timing. As Svensson and Woodford emphasize (e.g., Woodford 2005), it is not the overnight or one-month interest rate that mainly affects the economy, but the longer-term expected time path of interest rates. Surely any such surprise will affect future expected interest rates. The Bank forecasters will have to build into their forecasts some market reaction to that surprise, in order to guide the MPC as to whether enough has been done.

As Woodford (2005) notes:

> Another problem with the current procedure of the Bank of England is that it is unclear how the MPC is intended to determine the correct current repo rate in the event that the interest-rate path expected by the markets is judged to imply projections inconsistent with the Bank's target criterion. Would an attempt be made to determine the current repo rate that would lead to an acceptable projection, under the assumption that *the path of the repo rate after the current month would follow the path anticipated by the markets?* This would typically require an extreme adjustment of the current repo rate, as a change in the repo rate for only one month would have to change the path of inflation over the following two years by enough to get the projected inflation rate two years in the future on track. A more sensible approach would surely involve adjusting the entire path of interest rates to one that the MPC would view as more sound, rather than acting as if the committee expected itself to behave in the future in the way currently anticipated by the markets, even though it was planning to depart substantially from the markets' expectation in the short run. But in this case, projections would have to be produced on the basis of an assumption about future policy other than the one corresponding to market expectations. The idea that the MPC would be able to avoid taking a stand (at least in its internal deliberations) on a reasonable future path of interest rates, by insisting on using the markets' forecast in its projections, is not tenable.

Most often, however, in practice markets can, and do, anticipate *current* policy decisions reasonably well (see especially Lildholdt and Wetherilt 2004 for the United Kingdom). So this concern may be viewed as largely hypothetical. Moreover, if the problem was perceived as serious, then it could be largely met by also publicly revealing the adjustments made by the forecasters to the money-market yield curve to take account of estimated reactions.

Alternatively, and even simpler, since the inflation forecast is not published for a number of days after the MPC decision has been made, the forecasters could base their ex post forecast on the ex post reactions of the market to that decision. Admittedly, the choice of date(s) at which to measure the ex post reaction would be arbitrary, but then so too is the choice of dates on which to estimate the ex ante future path of rates. Moreover, should the market's reaction *not* be what the Bank/MPC wanted or expected, then the same argument as before—that the resulting published deviation of inflation from target should help to guide the market's expectation revisions—should presumably hold.

Even if the forecasters made *no adjustments* to take account of the current "surprise" decisions, so long as that was publicly known, then the published time path of inflation in the Inflation Report would give the market some idea of how the Bank expected that they *should* adjust their expectations; that is, if the current decision, followed by an unchanged path of future interest rates, led to inflation overshooting the target in the Inflation Report, then the market would be being guided to revise upward its expected future time path for interest rates.

A current concern is that few commentators seem to understand exactly on what basis the money-market yield curve used in the Bank of England's Inflation Report forecast *has* been constructed. Indeed, I have been led to understand that the ex ante forecast, *unadjusted* for the surprise element in the interest rate decision, continues to be used. This is reasonable so long as the surprise in the decision was minor, but what if it was not? Perhaps on such an occasion, the Bank/MPC would give some additional guidance.

But, in any case, and as noted earlier, there are limits to the extent of such "guidance" that the Bank of England can give by publishing a future deviation of inflation from target. In particular, a combination of a current surprise rise (fall) in the policy rate (perhaps to influence a current asset price boom or bust), together with a future forecast (mean) undershoot (overshoot) of inflation from target might be hard (but not impossible) to justify to the general public. It would probably be much harder to justify a surprise rise to offset an asset boom than a cut during a bust, as events in the second half of 2007 indicate. However, the question of whether the authorities respond asymmetrically to asset-price fluctuations (up and down), and whether this may matter, is outside the scope of this paper.

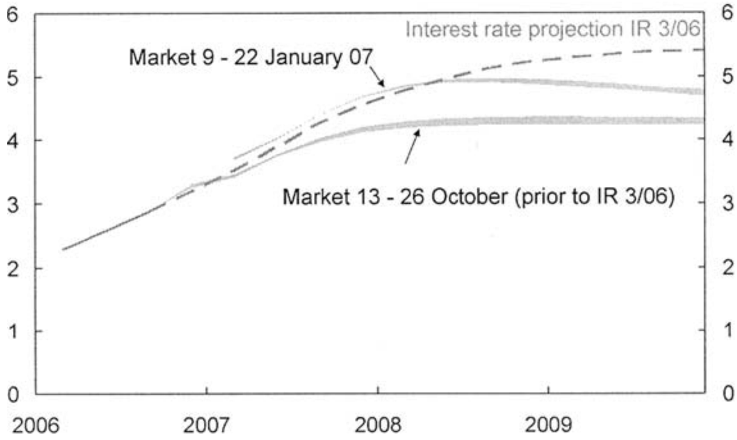
Just how serious these potential problems might ever become or—if they were perceived as serious—what steps might be taken in mitigation, is an issue that is beyond the scope or competence of this note. My gut feeling is that they probably would not be that serious in practice, but it does need careful watching. Be that as it may, I hope to have demonstrated that the UK MPC's current procedures on this front are not without their own inherent problems.

There are, also, somewhat similar problems with the use of a specific conditional policy forecast. How should the forecasters, for example, respond if the implied market yield curve does *not* then immediately move into line with the forecast set out by the MPC? The working assumption that is usually made is that the money-market yield curve will exactly, indeed slavishly, adjust to the MPC's prognostications. But this need not be so. Indeed, such a deviation is documented in a chart produced by Deputy Governor Bergo in a speech presented at the Foreign Exchange Seminar of the Association of Norwegian Economists, at which I was present (see Bergo 2007). This is shown as figure 3. When the Norges Bank interest rate projection of autumn 2006 was published, very short-term market forward interest rates did fall into line, but longer ones *did not.*[10] Another nice issue that has arisen

[10] The Deputy Governor noted the following:

> It is now almost three months since the previous Inflation Report was published. Since that time forward rates have increased and approached Norges Bank's interest rate path. Forward rates somewhat further out are still lower than our forecast. The reason may be that market participants have a different perception of the interest rate path that is necessary to stabilise inflation at target and to achieve stable developments in output and employment. Alternatively, the market may have the same short-term interest rate expectations as Norges Bank, but because of extraordinary conditions long-term bond prices are being pushed up and, consequently, long-term bond yields are being pushed down.

Figure 3. Norges Bank's Interest Rate Projection and Forward Rates



Source: Norges Bank.

in Norway is whether the Norges Bank is being time consistent in its own policy projections. This is addressed separately in the appendix.

There are questions about what such a discrepancy might imply and also how, if at all, it should be fed back into the next forecast. Should the forecasters give zero weight to the market (which, after all, now has the Norges Bank's prior policy forecast in its own information set and therefore has as much, or more, information than the MPC)? And, if not zero weight, what weighting in the MPC's forecast should be given to the discrepant forecasts?[11]

[11]This presumably depends on relative forecasting ability. That is dire, both for the central bank (see the chart in the *Financial Times*, January 30, 2007, on the NZ record) and for the market (for the United States, see Carriero, Favero, and Kaminska 2003, Diebold and Li 2003, Duffee 2002, Rudebusch 2002, and Rudebusch and Wu 2004; for Japan, see Thornton 2004). Wen Ben Lim and I intend to do further work on this for the United Kingdom. Perhaps for horizons longer than two quarters ahead, the constant interest rate assumption is not too bad after all.

Perhaps what the adoption of specific policy forecasts will do is to put more clearly under the academic microscope the (implicit or explicit) nature of the MPC's objective function and its time consistency. Academics will surely enjoy that exercise, but whether central bankers would also find that enjoyable is quite another question.

6. Conclusion

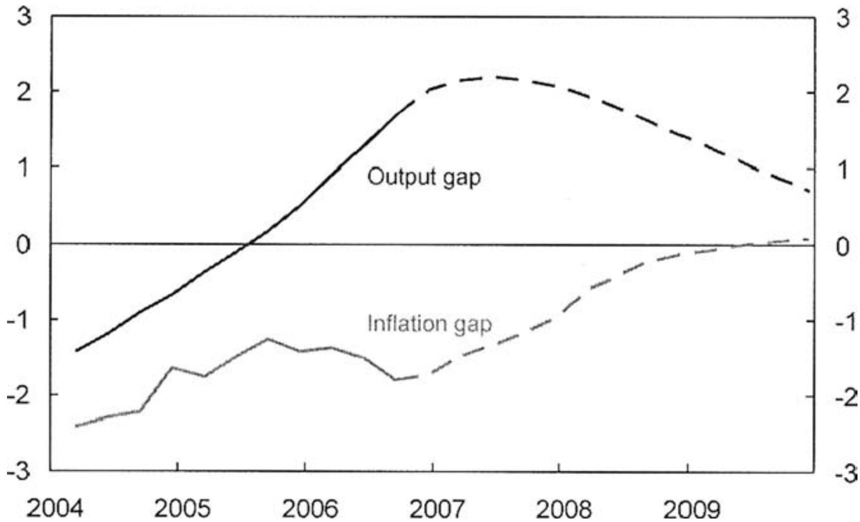
The constant interest rate (CIR) assumption had several beneficial aspects, one of which is an implicit humility about forecasting capabilities (official or market). But, under the influence of the recession of 2001–02, interest rates moved to such an exceptionally low level in many countries that the only plausible forecast/expectation was that they would revert to a higher, more normal level. The discrepancy between the latter plausible expectation and the CIR effectively led to the latter becoming untenable.

So what we now have, for those MPCs that reveal the basis of their conditional forecasts, is a choice between a market-based forecast and a forecast specifically chosen by the MPC. In *both* cases there will be problems of how to deal with discrepancies between these two alternatives. The specific forecast of the authorities should be (slightly) more informative, but there are offsetting problems. These problems include how to reach agreement in a committee of equals and whether the perception by the private sector of the extent of commitment of the MPC to its forecast path is properly aligned. Either way, what is fundamentally needed is a careful and candid description in accompanying statements and inflation reports of the thinking of each MPC. A picture (or graph) may paint a thousand words, but even such pictures need supporting explanations.

Appendix

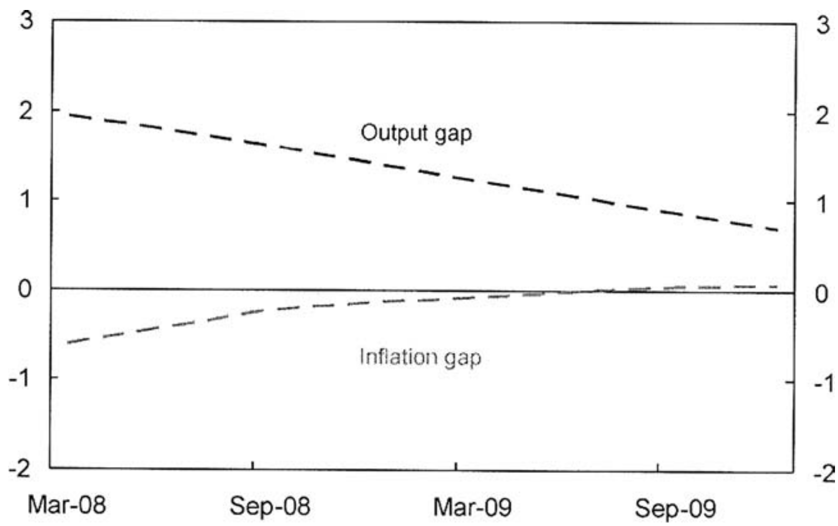
Consider the time paths for output and inflation produced in the Norges Bank forecast (March 2006), shown in figure 4, and then, assuming no shocks, just roll that same forecast forward to 2008 and 2009 (figure 5) (figures taken from the Deputy Governor's speech). In later years inflation is at target, but the output gap is still positive.

Figure 4. Trade-Off in Inflation Report, March 2006



Source: Norges Bank.

Figure 5. Trade-Off in Inflation Report, March 2006 2008–09



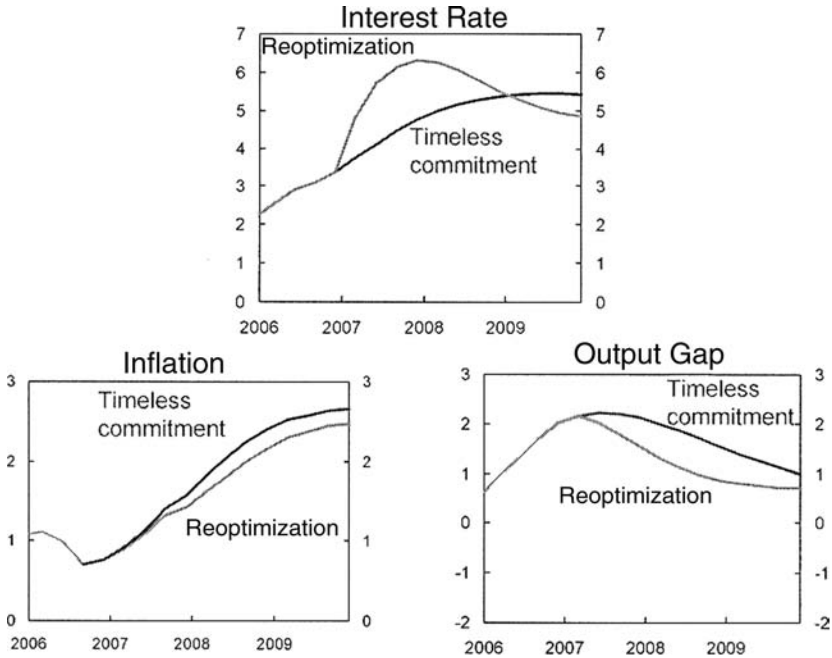
Source: Norges Bank.

If the loss function contains the output gap as an argument, this implies a time-varying coefficient upon it. The Deputy Governor commented as follows:

> Let us now take a closer look at our projections in the previous Inflation Report. The inflation gap closes gradually from below, while the output gap closes from above. According to the Bank's view, these paths provide a reasonable trade-off between the objective of stabilising inflation at target and stabilising developments in output and employment.
>
> Let us now use a time machine and travel forward to 2008. This picture, which is the same picture as the previous one but for a shorter time period, gives an impression that we place less weight on the output gap. The picture becomes even clearer if we travel forward yet another year in time to 2009.
>
> Inflation is now very near the target, while the output gap is still clearly positive. It may thus seem as if we are placing more weight on the output gap in the beginning of the period than at the end of the period. This suggests that the reference path in Inflation Report 3/06 is not consistent with a discretionary policy, where you make the best out of the situation in each period. Such a strategy would have involved a higher interest rate in order to provide a better balance between inflation and output towards the end of the projection period. Rather, it seems that the reference path has elements of commitment.
>
> Let us therefore assume that we follow the response pattern we have committed ourselves to earlier. In the literature, one such strategy is referred to as commitment under a timeless perspective.[12] It is possible to calculate, within the confines of our models, an optimal interest rate path based on such a strategy.
>
> In this example, we have been able to reconstruct (approximately) the reference path in Inflation Report 3/06 by minimising a loss function under commitment in a timeless perspective.

[12]See, for example, Woodford (1999). "Commentary: How Should Monetary Policy Be Conducted in an Era of Price Stability?" Paper presented at the Jackson Hole Conference, see http://www.columbia.edu/%7Emw2230/jhole.pdf.

Figure 6. Timeless Commitment versus Reoptimization



Source: Norges Bank.

> To reconstruct the reference path, the weight on the output gap in the loss function, lambda, has been set at 0.3. We also had to place a weight on changes in the interest rate in the loss function. This weight, which penalises large changes in the interest rate, can be defended based on considerations regarding robustness and financial stability.

That all sounds splendid, and academically very à la mode. The problem is that the alternative path of reoptimization (without commitment) using the same loss function, shown in figure 6, is extremely implausible. Would any central banker introduce a sharp, temporary spike in interest rates (in this case virtually doubling them), just to get output *lower* more quickly, and without that having much effect on getting, and keeping, inflation back to target?

References

Adolfson, M., S. Laséen, J. Lindé, and M. Villani. 2005. "Are Constant Interest Rate Forecasts Modest Interventions? Evidence from an Estimated Open Economy DSGE Model of the Euro Area." Sveriges Riksbank Working Paper No. 180 (March).

Berg, C., P. Jansson, and A. Vredin. 2004. "How Useful Are Simple Rules for Monetary Policy? The Swedish Experience." Sveriges Riksbank Working Paper No. 169.

Bergo, J. 2007. "Interest Rate Projections in Theory and Practice." Speech at the Foreign Exchange Seminar of the Association of Norwegian Economists at Sanderstølen, Norway, January 26.

Boivin, J. 2004. "Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data." Work in Progress (January version), Graduate School of Business, Columbia University.

Carriero, A., C. A. Favero, and I. Kaminska. 2003. "Financial Factors, Macroeconomic Information and the Expectations Theory of the Term Structure of Interest Rates." Unpublished Manuscript.

Diebold, F. X., and C. Li. 2003. "Forecasting the Term Structure of Government Bond Yields." NBER Working Paper No. 10048 (October).

Duffee, G. R. 2002. "Term Premia and Interest Rate Forecasts in Affine Models." *Journal of Finance* 57 (1): 405–43.

Edey, M., and A. Stone. 2004. "A Perspective on Monetary Policy Transparency and Communication." In *The Future of Inflation Targeting*, ed. C. Kent and S. Guttman, 73–100. (Proceedings of a conference held by the Reserve Bank of Australia, August 9–10.) Sydney: Reserve Bank of Australia.

Ehrmann, M., and M. Fratzscher. 2007. "Transparency, Disclosure, and the Federal Reserve." *International Journal of Central Banking* 3 (1): 179–225.

Goodhart, C. A. E. 2001a. "Monetary Transmission Lags and the Formulation of the Policy Decision on Interest Rates." *Review* (Federal Reserve Bank of St. Louis) 83 (4): 165–81.

———. 2001b. "The Inflation Forecast." *National Institute Economic Review* 175 (1): 59–66.

———. 2004. "What Is the Monetary Policy Committee Attempting to Achieve?" In *Macroeconomics, Monetary Policy, and Financial Stability: A Festschrift in Honour of Charles Freedman*. (Proceedings of a conference held by the Bank of Canada, June 2003.) Ottawa: Bank of Canada.

———. 2005. "The Monetary Policy Committee's Reaction Function: An Exercise in Estimation." *Topics in Macroeconomics* 5 (1): Article 18. Available at http://www.bepress.com/bejm/topics/vol5/iss1/art18.

Gürkaynak, R. S., B. Sack, and E. T. Swanson. 2005. "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements." *International Journal of Central Banking* 1 (1): 55–94.

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.

Jansson, P., and A. Vredin. 2003. "Forecast-Based Monetary Policy: The Case of Sweden." *International Finance* 6 (3): 349–80.

Lildholdt, P., and A. V. Wetherilt. 2004. "Anticipation of Monetary Policy in UK Financial Markets." Bank of England Working Paper No. 241.

Lomax, R. 2005. "Inflation Targeting in Practice: Models, Forecasts and Hunches." Speech at the 59th International Atlantic Economic Conference, London, March 12. Available in Bank of England's *Quarterly Bulletin* (Summer 2005).

———. 2007. "The MPC Comes of Age." Lecture at De Montfort University, Leicester, United Kingdom, February 28. Available at http://www.bankofengland.co.uk/publications/news/2007/022.htm.

Mishkin, F. S. 2004. "Can Central Bank Transparency Go Too Far?" In *The Future of Inflation Targeting,* ed. C. Kent and S. Guttmann, 48–66. (Proceedings of a conference held by the Reserve Bank of Australia, August 9–10.) Sydney: Reserve Bank of Australia.

Morris, S., and H. Shin. 1998. "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks." *American Economic Review* 88 (3): 587–97.

———. 2002. "Social Value of Public Information." *American Economic Review* 52 (5): 1521–34.

———. 2004. “Coordination Risk and the Price of Debt.” *European Economic Review* 48: 133–53.

Morris, S., H. Shin, and H. Tong. 2005. “Response to [L. Svensson’s] ‘Social Value of Public Information: Morris and Shin (2002) Is Actually Pro Transparency, Not Con’.” Paper in Progress (May).

Reifschneider, D. L., D. J. Stockton, and D. W. Wilcox. 1997. “Econometric Models and the Monetary Policy Process.” *Carnegie Rochester Conference Series on Public Policy* 47 (1): 1–37.

Romer, C. D., and D. H. Romer. 2004. “A New Measure of Monetary Shocks: Derivation and Implications.” Work in Progress (March version), University of California, Berkeley.

Rudebusch, G. D. 2002. “Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia.” *Journal of Monetary Economics* 49 (6): 1161–87.

Rudebusch, G. D., and T. Wu. 2004. “A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy.” Paper presented at the Federal Reserve Bank of San Francisco Conference on Interest Rates and Monetary Policy, San Francisco, CA, March 19–20.

Svensson, L. E. O. 2001. “Independent Review of the Operation of Monetary Policy in New Zealand.” Report to the Minister of Finance. Available at http://www.princeton.edu/∼svensson/.

———. 2003. “The Inflation Forecast and the Loss Function.” In *Central Banking, Monetary Theory and Practice*, ed. P. Mizen, 135–52. Cheltenham, UK: Edward Elgar.

———. 2004. “Monetary Policy with Judgement: Forecast Targeting.” Paper presented at the European Central Bank Conference on Monetary Policy and Imperfect Knowledge, Wurzburg, Germany, October 14–15.

———. 2005. “Social Value of Public Information: Morris and Shin (2002) Is Actually Pro Transparency, Not Con.” NBER Working Paper No. W11537 (August).

Thornton, D. 2004. “Testing the Expectations Hypothesis: Some New Evidence for Japan.” Discussion Paper No. 2004-E-3 (February), Bank of Japan, Institute for Monetary and Economic Studies.

———. 2003. *Interest & Prices*. Princeton, NJ: Princeton University Press.

———. 2004. "Inflation Targeting and Optimal Monetary Policy." *Review* (Federal Reserve Bank of St. Louis) 86 (4): 15–41.

———. 2005. "Central Bank Communication and Policy Effectiveness." Paper presented at the Sveriges Riksbank Conference on Inflation Targeting, Stockholm, Sweden, June 10–12.

Futures Contract Rates as Monetary Policy Forecasts*

Giuseppe Ferrero and Andrea Nobili
Bank of Italy

The prices of futures contracts on short-term interest rates are commonly used by central banks to gauge market expectations concerning monetary policy decisions. Excess returns—the difference between futures rates and the realized rates—are positive, on average, and statistically significant, both in the euro area and in the United States. We find that these biases are significantly related to the business cycle only in the United States. Moreover, the sign and the significance of the estimated relationships with business-cycle indicators are unstable over time. Breaking the excess returns down into risk-premium and forecast-error components, we find that risk premia are countercyclical in both areas. On the contrary, ex post prediction errors, which represent the greater part of excess returns at longer horizons in both areas, are negatively correlated with the business cycle only in the United States.

JEL Codes: E43, E44, E52.

1. Introduction

In order to infer market expectations about the future course of monetary policy, central banks commonly use prices of financial assets and survey data. The former are available at high frequencies, but they also incorporate risk and term premia, which may distort their

*We thank Paolo Angelini, Pier Paolo Benigno, and Alessandro Secchi. We are also grateful for helpful comments to seminar participants at the North American Summer Meeting of the Econometric Society in Minneapolis and at "The Analysis of the Money Market: Role, Challenges and Implications from the Monetary Policy Perspective" Conference organized by the European Central Bank. All remaining errors are ours. The views are personal and do not necessarily reflect those of the institution with which we are affiliated. Author e-mails: giuseppe.ferrero@bancaditalia.it and andrea.nobili@bancaditalia.it.

information content in terms of expected future interest rates, while the latter are likely not to be affected by premia but are available at a relatively low frequency. Both measures might be biased estimators of ex post realized interest rates to the extent that they incorporate systematic forecast errors.

Recent studies for the United States have compared the information content of several financial instruments, finding that yield curves and futures contracts on short-term interest rates are good predictors of the future path of monetary policy decisions both in the short and medium term (Piazzesi and Swanson 2004; Gürkaynak, Sack, and Swanson 2006). Nevertheless, another strand of the literature has provided evidence that ex post excess returns—namely, the differences between short-term interest rates implied in the price of Eurodollars futures and the ex post realized spot rates—are, on average, positive and statistically significant (Krueger and Kuttner 1996; Sack 2002; Durham 2003). Recently, Piazzesi and Swanson (2004) have shown that this bias is time varying, countercyclical, and predictable by means of business-cycle indicators. This finding suggests that policymakers should look at adjusted measures of futures rates in order to assess the efficacy of their communication more accurately.

The label "risk premia" is often used in the financial literature to refer to predictable excess returns on the short-term interest rate (Piazzesi and Swanson 2004; Cochrane 2006). However, risk premia and predictable excess returns do not necessarily coincide. For example, in the presence of structural breaks, economic agents may need time to learn about the new environment: in the early stages of this process, previously held beliefs could lead to a long series of errors all in the same direction until forecasters finally learn about the structural break. In this case ex post excess returns may incorporate two predictable components. One is the ex ante risk premium, defined as the difference between the futures rates and the market expectation of future spot interest rates, which is required by investors when they buy or sell the financial contract. The other is a systematic prediction error.

In this respect, this paper reassesses the predictive power of short-term interest rate futures by extending the analysis of Piazzesi and Swanson (2004) along two dimensions. First, we use futures contracts on short-term interest rates in euros and investigate the size and the magnitude of ex post excess returns in the euro area, allowing a comparison with those in the United States. Second, we rely on professional forecast surveys in order to disentangle the risk premium and forecast-error components of ex post excess returns and to study their behavior over the business cycle.

Our empirical investigation reveals that euro-area ex post excess returns are of the same sign and magnitude as those in the United States, but they do not appear to be significantly related to the business cycle. In addition, the relation between excess returns and the business cycle appears to be unstable over time in both areas. This evidence is in contrast with the findings of the recent strand of the literature that studies term-structure models, which suggests that the implied risk premia should be strongly affected by business-cycle fluctuations.

We show that these puzzling results essentially depend on the common assumption that ex post excess returns coincide entirely with risk premia. Our proposed empirical breakdown of ex post excess returns suggests that risk premia are, on average, not significantly different in the United States and in the euro area, and are significantly countercyclical in both areas. Interestingly, the predictive regressions involving risk premia and business-cycle indicators are stable over time. By contrast, ex post prediction errors, which represent the largest fraction of the whole excess return at longer horizons in both areas, are significantly and negatively related to the business cycle only in the United States.

We argue that our excess returns decomposition has important implications for central banks when they assess financial markets' expectations regarding the future path of monetary policy decisions. Even though interest rate futures adjusted for both components provide the best forecast of future spot interest rates, they no longer coincide with financial markets' view. Policymakers should assess markets' expectations about future interest rates by looking at quoted futures rates adjusted by the premia component only, as the ex post prediction error reflects part of the expectations formation process.

The remainder of the paper is organized as follows. In section 2 we describe the data set used in the analysis. In section 3 we provide evidence on the size and predictability of ex post excess returns on short-term interest rates in euros, allowing a comparison with those in dollars. In section 4 we decompose ex post realized excess returns into risk premia and systematic prediction errors and investigate their relation with the business cycle. In section 5 we point out the main implications of our proposed breakdown for policymakers. Section 6 concludes.

2. The Data Set

We define the ex post excess return realized from holding the n-quarter-ahead contract to maturity as

$$x_{t+n}^{(n)} = f_t^{(n)} - r_{t+n}, \tag{1}$$

where $f_t^{(n)}$ denotes the average of the futures contract rates quoted on the first ten days of the last month of quarter t for a contract expiring at the end of quarter $t + n$ and r_{t+n} is the corresponding realized spot interest rate prevailing on the day of expiration of the futures contract.[1]

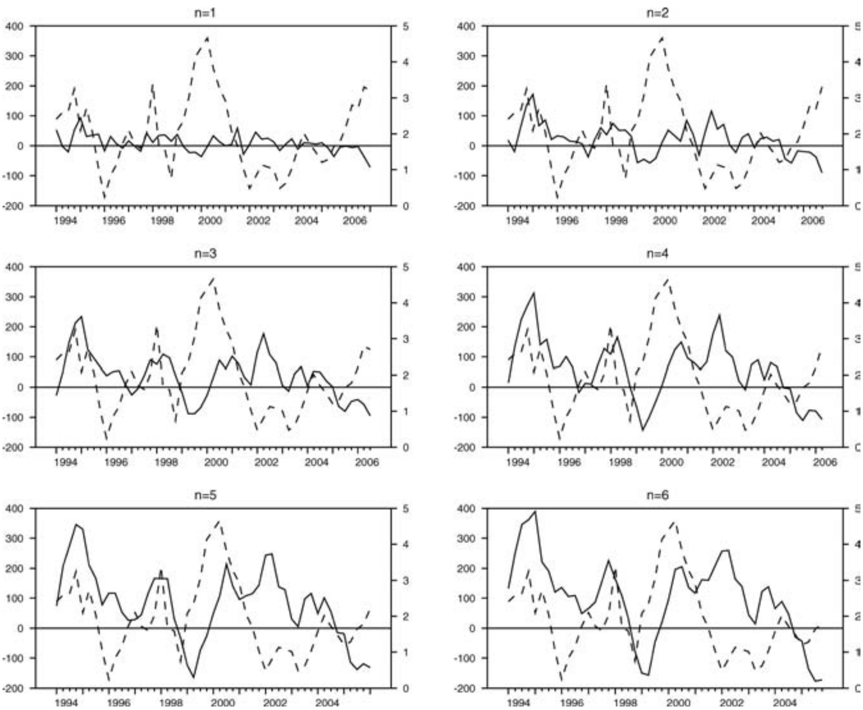
Regarding the euro area, we restrict our attention to futures contracts on short-term interest rates traded on the London International Financial Futures Exchange (LIFFE), which mature two business days prior to the third Wednesday of the delivery month. At each point in time we focus on the first six (unexpired) contracts.[2] The choice of the sample period, 1994–2007, reflects the limited availability of survey data used for the excess returns decomposition, which is the core of our analysis. In particular, for the pre-EMU period (1994:Q1–1998:Q4), we consider futures contracts linked to the British Bankers' Association offered rate (BBA LIBOR) for three-month Eurodeutschmark deposits. The idea is that the institutional features and anti-inflationary objective of the European Central Bank's (ECB's) monetary policy largely resemble those of the German Bundesbank.[3] For the EMU period (1999:Q1–2007:Q1) we

[1] Results do not change significantly using the futures contract rate quoted on the last trading day of quarter t.

[2] By far, the most actively traded futures contracts on three-month deposits are those with delivery in March, June, September, and December.

[3] Buiter (1999) suggests that the ECB adheres to a "priestly" view of central banking in that it adopts "many of the procedures and practices of the old Bundesbank."

Figure 1. Ex Post Excess Returns (Solid Line) and Real GDP Growth (Dashed Line) in the Euro Area



Notes: The sample period is 1992:Q1–2007:Q1. Ex post excess returns are measured in basis points.

focus on contracts whose underlying asset is the European Banking Federation's Euribor Offered Rate (EBF Euribor) for three-month euro deposits. For the United States we compute the ex post excess returns using futures contracts on three-month LIBOR Eurodollar deposit rates, which are quoted on the Chicago Mercantile Exchange.

Figure 1 plots the time series of the ex post realized excess returns on futures contracts in euros expiring up to six quarters ahead. Three basic features emerge. First, independently from the forecasting horizon, these returns are generally positive, suggesting that futures rates are, on average, higher than ex post realized spot rates. Second, they increase with the forecast horizon, consistently with the view that agents demand larger term premia on contracts with longer expiration dates. Third, they move significantly over time (see also Piazzesi and Swanson 2004).

3. Reassessing Ex Post Excess Returns

3.1 Constant Excess Returns

We start our analysis by checking whether futures contract rates are unbiased predictors of spot short-term interest rates. To this end, we follow Piazzesi and Swanson (2004) and regress the computed ex post excess returns on a constant term

$$x_{t+n}^{(n)} = \alpha^{(n)} + \epsilon_{t+n}^{(n)} \tag{2}$$

for the forecast horizons $n = 1, 2, 3, \ldots, 6$ quarters and test in each equation whether the estimated coefficients $\alpha^{(n)}$ are different from zero.

In the absence of arbitrage opportunities, this analysis is also considered a test of the validity of the (pure) rational-expectations hypothesis—namely, that futures contract rates are, on average, equal to the expected spot interest rates.[4] We notice that in the financial literature (Fama 1984; Campbell and Shiller 1991; Campbell 1995) the validity of this hypothesis has also been tested by running predictive regressions of the type

$$r_{t+n}^{(n)} = \alpha^{(n)} + \beta^{(n)} f_t^{(n)} + \epsilon_{t+n}^{(n)} \tag{3}$$

and performing the joint test of the null hypothesis that $\alpha^{(n)} = 0$ (zero-mean term premia) and $\beta^{(n)} = 1$ (no time-varying term premia).[5] However, some drawbacks of this second approach have been recently stressed. First of all, standard errors in regressions of this

[4] In the weaker version of the forward-rate expectation hypothesis, the constant term is allowed to be nonzero.

[5] Interestingly, Gürkaynak, Sack, and Swanson (2006) find that the hypothesis that $\beta = 1$ cannot be rejected for a number of U.S. market interest rates. This evidence, they say, suggests only that the time-varying excess returns are not correlated enough with the ex post spot interest rates spreads to drive the estimated coefficients far from one. It does not rule out the possibility that they are correlated with other variables, such as business-cycle indicators.

type are typically large enough that the expectations hypothesis cannot be rejected, as regression tests are not powerful enough to distinguish between the expectations hypothesis and alternative hypotheses in a sample of the length considered here (Kim and Orphanides 2005). Moreover, equation (3) may raise concerns regarding spurious correlation among variables, insofar as spot interest rates and futures contract rates are nonstationary variables. Although the results could be strongly sample dependent, there is some evidence that various international nominal short- and long-term interest rates may contain a unit root in the levels of the series (e.g., Rose 1988; Rapach and Weber 2004).[6]

Results for the estimated coefficients of equation (2) are summarized in table 1, where standard errors are computed by means of the Newey-West heteroskedasticity and autocorrelation-consistent procedure, in order to take into account the futures contracts overlapping. In the euro area the average ex post realized excess returns are significantly positive over the sample period, ranging from about 10 basis points at the one-quarter horizon to 100 basis points at the six-quarter horizon.

A corresponding analysis for the United States suggests that ex post excess returns have likewise been significantly positive and also slightly larger than those obtained for the euro area, ranging from about 20 basis points at the one-quarter horizon to 110 basis points at the six-quarter horizon.[7]

3.2 Time-Varying Excess Returns

Relying on previous studies for the U.S. Treasury market (Fama and Bliss 1987; Cochrane and Piazzesi 2002) and, more recently, for quoted futures rates (Piazzesi and Swanson 2004), we assess whether

[6]In order to deal with nonstationarity, the validity of the expectations hypothesis is usually tested by subtracting the current level of spot rates or first-differencing the variables in equation (3) (Jongen, Verschoor, and Wolff 2005; Gürkaynak, Sack, and Swanson 2006).

[7]In annualized terms, excess returns in the euro area range from 34 basis points at the one-quarter horizon to 68 basis points at the six-quarter horizon; in the United States they range from 73 to 75 basis points. In the sample period 1985:Q1–2005:Q4, Piazzesi and Swanson (2004) find that the average annualized excess returns range from 60 basis points at the one-quarter horizon to 100 basis points at the six-quarter horizon.

Table 1. Constant Excess Returns

Euro Area						
n	1	2	3	4	5	6
$\alpha^{(n)}$	8.4** (4.4)	20.5** (9.8)	37.7** (16.8)	59.1** (23.3)	80.7** (28.9)	102.2** (32.9)
United States						
$\alpha^{(n)}$	18.3** (6.0)	33.3** (14.7)	51.7* (25.5)	73.6** (34.5)	93.6** (42.8)	112.2** (49.6)

Notes: The sample period is 1994:Q1–2007:Q1. Ex post excess returns are measured in basis points. Predictive regressions are estimated by OLS. Newey-West standard errors are reported in parentheses. * denotes significance at the 10 percent confidence level; ** denotes significance at the 5 percent level.

the term structure of interest rates implied in futures contracts in euros is also characterized by time-varying and predictable excess returns. The predictability of excess returns is explored by running the following regressions,

$$x_{t+n}^{(n)} = \alpha^{(n)} + \beta^{(n)} z_t + \gamma^{(n)} f_t^{(n)} + \epsilon_{t+n}^{(n)}, \tag{4}$$

which involve a business-cycle indicator observable at time t—namely, z_t—and the level of the futures rate itself. Under the assumption that excess returns can be interpreted as risk premia, their predictability using business-cycle indicators finds theoretical foundation in standard asset pricing models (Cochrane 2006), while the broader specification in (4), which includes the futures rate as an additional regressor, essentially relies on the recent strand of the financial literature that uses the affine structure to model the yield curve and the price of risk. These studies typically employ Gaussian affine term-structure models in which time-varying risk premia depend on two latent factors usually identified, respectively, with the level of the short-term interest rate and the slope of the yield curve. The significant relationship between the yield curve and observable state variables reflecting business-cycle fluctuations has been amply documented in Ang and Piazzesi (2003), Ang, Dong, and Piazzesi (2004), Rudebusch and Wu (2004), Ang, Piazzesi, and

Wei (2006), Hördal, Tristani, and Vestin (2006), and Pericoli and Taboga (2006).[8]

Results for the euro area are reported in the top part of table 2 and refer to two business-cycle indicators. For each maturity, the first column shows the estimated coefficients obtained using the annual growth rate of real GDP, which is commonly considered the most natural proxy for the business cycle. Because official real GDP data are released with a lag and frequently revised, there may be significant differences between the data used in the regression and the data available to market participants at the time contract prices were settled. To avoid this problem, we perform real-time predictive regressions using real GDP lagged one quarter and alternative business-cycle indicators. In particular, we use indices from the European Commission's survey of manufacturing industry, household consumption, construction, and retail trade. In order to select a narrower set of variables from the large volume of available survey data, we performed a preliminary cross-correlation analysis at business-cycle frequencies between each of them and real GDP. Among the variables with greater contemporaneous correlation, we find that "employment expectations for the months ahead" in manufacturing industry has the best properties in terms of significance and goodness of fit in regression (4).[9] As the survey is available at monthly frequency, in our quarterly regressions we include the data for the second month of the quarter considered, in order to avoid the use of data not available when agents form their expectations. Moreover, in order to compare the results obtained with different variables and between the two areas, we normalize the regressors to have zero mean and unit variance. Excess returns on futures contracts in euros do not appear to be significantly related to the business cycle.

Table 2 allows us to compare the predictability of excess returns in the two areas in the same sample period. For the United States, we

[8]For a survey, see Diebold, Piazzesi, and Rudebusch (2005).

[9]The contemporaneous correlation of this variable with real GDP at business-cycle frequencies is 0.6. We also run regressions including simultaneously two or more business-cycle indicators and involving one or more estimated common factors obtained from a dynamic factor model based on all the considered business-cycle indicators. Results in terms of goodness of fit are not better than those obtained with employment expectations. The results obtained with other survey data are available from the authors upon request.

Table 2. Time-Varying Excess Returns

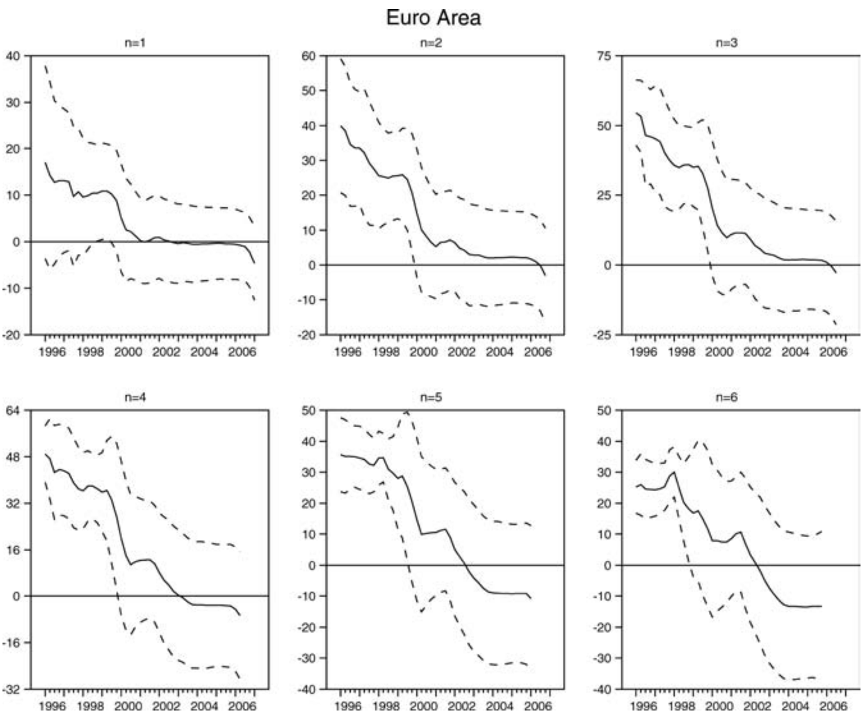
Euro Area	$n = 1$		$n = 2$		$n = 3$		$n = 4$		$n = 5$		$n = 6$	
Constant	8.4**	8.4**	20.5**	20.5**	37.7**	37.7**	59.1**	59.1**	80.7**	80.7**	102.2**	102.2**
	(3.8)	(4.1)	(8.1)	(8.7)	(13.0)	(14.0)	(16.3)	(17.3)	(18.5)	(19.2)	(19.8)	(20.2)
RGDP	−13.5**	—	−18.2	—	−20.0	—	−22.5	—	−22.5	—	−20.8	—
	(4.1)		(11.1)		(17.6)		(18.5)		(16.7)		(14.8)	
E(empl)	—	−4.7	—	−3.0	—	−2.8	—	−6.7	—	−10.7	—	−13.2
		(4.6)		(9.2)		(13.9)		(15.3)		(15.1)		(14.3)
Future	16.7**	10.4**	32.6**	22.9**	52.2**	41.8**	79.0**	68.3**	100.9**	91.4**	115.9**	108.0**
	(4.9)	(4.6)	(13.0)	(9.6)	(17.5)	(12.7)	(17.5)	(13.7)	(16.7)	(14.4)	(16.8)	(15.8)
R^2	0.22	0.09	0.24	0.16	0.32	0.27	0.46	0.43	0.57	0.55	0.63	0.62
United States												
Constant	18.3**	18.3**	33.3**	33.3**	51.7**	51.7**	73.6**	73.6**	93.6**	93.6**	112.2**	112.2**
	(5.6)	(5.7)	(12.9)	(12.9)	(21.4)	(19.2)	(28.0)	(21.8)	(33.4)	(22.8)	(37.3)	(24.2)
RGDP	−14.7*	—	−35.2**	—	−57.9**	—	−76.5**	—	−91.5**	—	−103.2**	—
	(8.2)		(14.1)		(20.8)		(25.7)		(30.5)		(32.9)	
NFP	—	−22.8**	—	−73.8**	—	−127.2**	—	−177.4**	—	−211.8**	—	−224.6**
		(10.6)		(23.0)		(28.9)		(27.2)		(24.6)		(24.9)
Future	11.6*	22.8**	30.8**	74.5**	54.7**	131.2**	81.5**	189.5**	109.2**	234.9**	134.6**	261.5**
	(6.8)	(10.7)	(12.1)	(23.4)	(18.1)	(28.8)	(24.1)	(30.1)	(28.5)	(31.1)	(32.9)	(32.9)
R^2	0.06	0.04	0.12	0.19	0.18	0.33	0.24	0.48	0.31	0.61	0.37	0.66

Notes: The sample period is 1994:Q1–2007:Q1. RGDP is real GDP growth rate. E(empl) is employment expectations for the months ahead from the industrial survey by the European Commission. NFP is the growth rate of nonfarm payrolls. Ex post excess returns are measured in basis points. All predictive variables are standardized. Predictive regressions are estimated by OLS. Newey-West standard errors are reported in parentheses. * denotes significance at the 10 percent confidence level; ** denotes significance at the 5 percent confidence level.

use as business-cycle indicators the annual growth of real GDP and the real-time year-on-year change in nonfarm payrolls. In this case, our estimates confirm the results obtained by Piazzesi and Swanson (2004) for the sample period 1985–2005. The slope coefficients are, in general, highly significant and negative, and their size increases with the forecast horizon. However, some concerns may arise with these estimates.

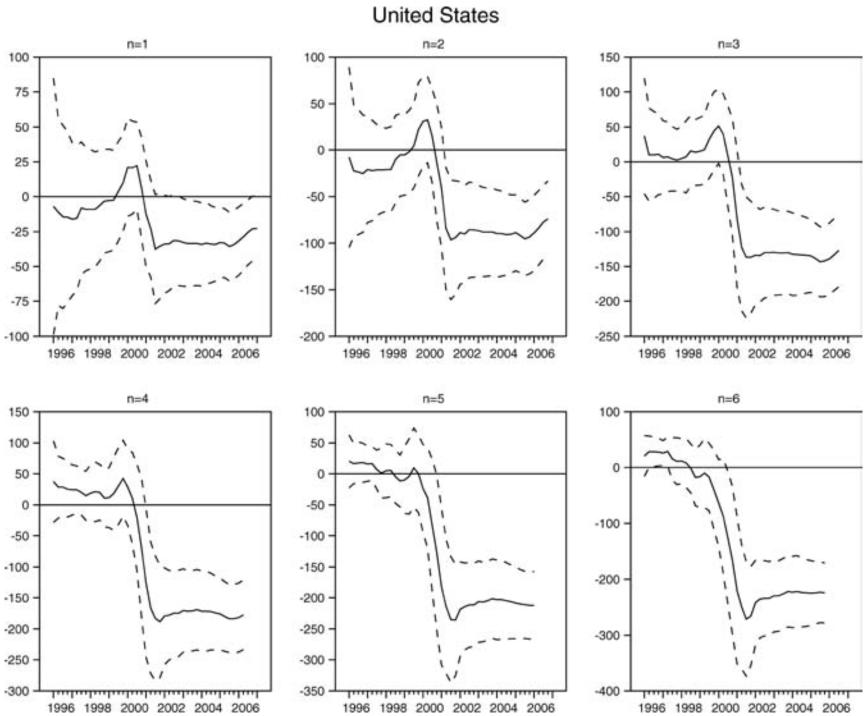
A first issue is the stability of the estimated coefficients. In figure 2 we plot the recursive estimates of coefficients of the business-cycle indicator used in equation (4). Interestingly, the coefficients decreased significantly over time both in the euro area and in the United States. In particular, we cannot exclude that the coefficients

Figure 2. Recursive Coefficients for the Business-Cycle Indicator



(continued)

Figure 2. (Continued)



Notes: Recursive least-squares estimates. The initial estimate is obtained using the sample 1994:Q1–1996:Q1. Employment expectations for the months ahead are used in predictive regressions for the euro area. Nonfarm payrolls are used in predictive regressions for the United States. Dotted lines represent the two standard-error bands around the estimated coefficients.

were positive in the period 1994–2000 and became negative afterwards. The cumulative sum (CUSUM) tests for overall stability of the estimated regressions show significant departures of the computed test statistics from their expected value, thus providing evidence for the presence of parameter or variance instability in the predictive regressions (figure A1 in the appendix).

Another important concern is that excess returns may be nonstationary in the sample period. To the extent that the regressor variables are also nonstationary, the interpretation of the previous estimated predictive regressions may prove erroneous. In table 3 we investigate the time-series properties of the variables used in the

Table 3. Unit-Root Test

	Euro Area		United States	
	No Trend	**Linear Trend**	**No Trend**	**Linear Trend**
$x_t^{(1)}$	−3.687**	−5.188**	−2.734**	−3.057*
$x_t^{(2)}$	−3.106**	−3.677**	−2.828**	−3.103*
$x_t^{(3)}$	−3.014**	−3.620**	−1.735*	−2.041
$x_t^{(4)}$	−2.613**	−2.983**	−1.723*	−1.928
$x_t^{(5)}$	−2.726**	−3.067**	−1.591	−1.704
$x_t^{(6)}$	−2.573**	−2.960**	−1.382	−1.914
$f_t^{(1)}$	−1.395	−2.027	−1.639	−2.568
$f_t^{(2)}$	−1.671	−2.368	−1.271	−1.411
$f_t^{(3)}$	−1.446	−1.881	−1.385	−1.630
$f_t^{(4)}$	−1.536	−2.098	−1.477	−1.848
$f_t^{(5)}$	−1.580	−2.307	−1.568	−2.077
$f_t^{(6)}$	−1.520	−2.410	−1.634	−2.287
z_t	−1.180	−2.346	−1.807	−2.109

Notes: The sample period is 1994:Q1–2007:Q1. The lag order p has been selected using a Schwarz information criterion with the maximum lag length of 8. * denotes the rejection of the null hypothesis at the 10 percent level; ** denotes the rejection of the null hypothesis at the 5 percent confidence level.

predictive regressions by means of the modified augmented Dickey-Fuller test (DF-GLS) for unit root (Elliot, Rothenberg, and Stock 1996).[10]

While excess returns on futures contracts in euros appear to be stationary at all maturities, for those in dollars we cannot reject the hypothesis that they contain a unit root, at least at horizons longer than two quarters. Strong evidence of nonstationarity is also found for futures rates in both areas, while for the business-cycle indicators the evidence is less clear-cut and needs to be treated with

[10]In order to discriminate whether the variables of interest are stationary around a deterministic trend, we also show the results by including in the test regression both the constant term and a linear trend.

caution because of the relatively low power of tests in small samples. These findings suggest that the significant relation between excess returns and the business cycle in the United States may simply reflect a common long-run trend but not short-run co-movements among variables.[11]

To the extent that we interpret excess returns as proxies for risk premia, the results of the previous predictive regressions are puzzling. Why in the overall sample do risk premia behave so differently in the two areas? Why has the relation between the business cycle and the risk premia changed over time?

4. Understanding Excess Returns: A Decomposition

First of all, we argue that the previously estimated regressions provide correct measures of the risk premia only under the crucial assumption that the agents are perfectly rational—namely, that they do not make systematic errors in their predictions.[12] In that case, prediction errors are orthogonal to the information set and the only predictable part of the excess return is the risk premium.

However, the financial literature suggests that prices may differ systematically (at least for a period of time) from what people expected them to be for different reasons: (i) prices reflect information to the point where the marginal benefits of acting on information do not exceed the marginal cost (Fama 1991); (ii) agents may rationally process only a limited amount of information because of capacity constraints (Sims 2003); (iii) even if forecasts are formed rationally, allowing for large interest rate movement with small probability, the forecast will appear biased when judged ex post (the so-called peso problem) (Bekaert, Hodrick, and Marshall 2001); (iv) in a changing environment, agents in the market form expectations

---

[11] We have also estimated the predictive regressions using techniques that take into account the nonstationarity of time series, such as dynamic OLS (e.g., Stock and Watson 1993), fully modified least squares (e.g., Phillips and Hansen 1990), and the vector error correction model (e.g., Johansen 1991, 1995). We find the long-run relationships between excess returns and predictive variables to be significant at horizons longer than one quarter.

[12] The concept of rational expectations, as described in Sargent (1986) asserts that outcomes should not differ systematically (i.e., regularly or predictably) from what people expected them to be.

by learning from past experience (Timmermann 1993) or they are subject to irrational exuberance (Shiller 2000).[13]

In all these cases, ex post excess returns realized from holding the n-quarter-ahead futures contract to maturity may embody two predictable components:

$$x_{t+n}^{(n)} = \theta_t^{(n)} + \sigma_{t+n}^{(n)}, \tag{5}$$

where

$$\theta_t^{(n)} = f_t^{(n)} - E(i_{t+n}|I_t) \tag{6}$$

and

$$\sigma_{t+n}^{(n)} = E(i_{t+n}|I_t) - i_{t+n}. \tag{7}$$

The first component, $\theta_t^{(n)}$, is the ex ante risk premium, defined as the difference between the futures rates and the market expectation of future spot interest rates, conditional on the information set available to the agents at time t. The second component, $\sigma_{t+n}^{(n)}$, is the ex post prediction error made by market participants in forecasting future spot rates and is measured as the difference between the conditional expectation on future interest rates and ex post realized spot rates. As in absence of perfect rationality this second component may be (at least in the short run) systematically different from zero, ex post excess returns can differ substantially from risk premia.

As a proxy for market expectations, $E(i_{t+n}|I_t)$, we consider the mean of short-term interest rate forecasts from the Consensus Forecast survey. This survey has the advantage of providing a long time series on a quarterly basis regarding expectations on future short-term interest rates at horizons up to eight quarters ahead.

The use of survey forecasts may raise concerns for several reasons. The most important one in our context is that, in principle, survey respondents may just use the unadjusted futures contract rates in

[13] There is growing empirical evidence, based mainly on survey data, that the perfect-rationality assumption is violated for expectations on many macroeconomic and financial variables and for many industrialized countries, including the United States and members of the EMU (e.g., Froot 1989; Gourinchas and Tornell 2004; Jongen, Verschoor, and Wolff 2005; Bacchetta, Mertens, and van Wincoop 2006).

order to provide their own forecasts on future spot short-term interest rates. In this case, the forecast would also incorporate the premia component, and the ex post forecast error would be observationally equivalent to the original excess return. Since most of the respondents to the Consensus Forecast survey are professional forecasters who work for institutions operating in the financial markets, even though they may differ from people operating directly in the market, it is likely that they share their information. Therefore, it seems reasonable to assume that respondents to the survey are able to separate the premium component from the forecast component. This hypothesis is also supported by evidence presented by Kim and Orphanides (2005) for the United States that shows that survey expectations on short-term interest rates based on the Blue Chip Financial Forecast incorporates the premium correction.

The estimates of the average value of the two components are obtained by running the regressions[14]

$$\sigma_{t+n}^{(n)} = \alpha_{\sigma}^{(n)} + \epsilon_{t+n}^{(n)} \tag{8}$$

$$\theta_{t}^{(n)} = \alpha_{\theta}^{(n)} + \eta_{t}^{(n)}. \tag{9}$$

Results are reported in table 4. The estimates show that in the euro area, average risk premia are significant at all forecast horizons and are smaller than the corresponding systematic forecast errors at horizons longer than two quarters. In particular, the ex ante risk premium ranges from about 10 to 35 basis points, while the systematic prediction error is between 0 and 70 basis points (see also figure A2 in the appendix).[15] The former

[14]Consensus Economics receives the answers of the survey the first Friday of the last month of the quarter in which it publishes the results of the survey. Since the risk premia are computed using the averages of the market prices of futures contracts quoted on the first ten trading days of the month in which the quarterly Consensus Forecast survey is published, the information sets of respondents to the Consensus Forecast survey and market operators should not be significantly different. In order to verify that the information sets of market participants are not too different, the predictive regressions have been also estimated using spot data from various days on either sides of the first Friday of the last month of the quarter. The results are robust to this modification.

[15]In annualized terms, the risk premia range from 36 to 23 basis points and prediction errors from about 0 to 45 basis points.

Table 4. Excess Returns Decomposition

Euro Area						
n	**1**	**2**	**3**	**4**	**5**	**6**
$\theta^{(n)}$	9.1** (2.3)	13.6** (4.5)	17.7** (6.3)	24.5** (8.2)	30.4** (9.5)	34.2** (10.6)
$\sigma^{(n)}$	−0.7 (4.4)	6.9 (9.8)	19.9 (16.1)	34.6 (21.4)	50.3** (25.6)	68.0** (29.5)
United States						
$\theta^{(n)}$	12.2** (3.8)	17.6** (5.4)	25.1** (6.6)	32.2** (7.2)	37.9** (8.4)	42.0** (9.0)
$\sigma^{(n)}$	6.0 (5.4)	15.7 (10.0)	26.6* (14.8)	41.5** (18.6)	55.7** (22.0)	70.2** (24.8)
Estimated Coefficients for Risk Premia (tbill3m–LIBOR3m)						
$\phi^{(n)}$	28.6** (2.6)	28.7** (2.7)	28.6** (2.7)	28.3** (2.7)	28.1** (2.8)	28.1** (2.8)
$\gamma^{(n)}$	10.9** (2.0)	11.1** (2.1)	11.3** (2.1)	11.7** (2.2)	11.9** (2.2)	12.2** (2.3)

Notes: The sample period is 1994:Q1–2007:Q1. $\theta_1^{(n)}$ and $\sigma_1^{(n)}$ refer to the subsample period 1994:Q1–1998:Q4; $\theta_2^{(n)}$ and $\sigma_2^{(n)}$ refer to the subsample period 1999:Q1–2007:Q1. Ex ante risk premia and ex post forecast errors are measured in basis points. Predictive regressions are estimated by OLS. Newey-West standard errors are reported in parentheses. * denotes significance at the 10 percent confidence level; ** denotes significance at the 5 percent confidence level.

component accounts for more than 60 percent of the overall predictable excess returns at the two-quarter horizon, for about 50 percent at the three-quarter horizon, and for about 40 percent at longer horizons.

For the United States, the Consensus Forecast survey reports expectations on the three-month Treasury-bill rate, which may differ from the three-month LIBOR because of the existence of different premia (Campbell and Shiller 1991; Cochrane and Piazzesi 2002). Therefore, the ex ante risk premium, $\widehat{\alpha}_{\sigma}^{(n)}$, is obtained by adjusting the Consensus Economics forecast for an estimated time-varying premium

$$PR_t \equiv i_t - tb_t = \phi + \tau x_t + e_t, \tag{10}$$

where i_t is the money-market rate (three-month LIBOR) and tb_t is the three-month Treasury-bill rate.[16] In table 4 we report the results of the nonlinear least-squares joint estimation of the two different premia:

$$\theta_t^{(n)} \equiv f_t^{(n)} - E_t[tb_{t+n}] - PR_t = \alpha_\sigma^{(n)} + \epsilon_{t+n}^{(n)} \tag{11}$$

$$PR_t = \phi^{(n)} + \tau^{(n)} x_t + e_t^{(n)}. \tag{12}$$

Average risk premia in the United States, $\theta_t^{(n)}$, range between 10 and 40 basis points; they are not significantly different from those in the euro area at all horizons and they account for about 50 percent of the overall excess return at the two-quarter and three-quarter horizons and for about 40 percent at longer horizons.[17] Systematic prediction errors started to increase significantly in 2000 (see figure A3 in the appendix), when the Federal Reserve stopped announcing its expected future policy stance ("policy bias"), and returned to the lowest level in 2003, when the Federal Open Market Committee reintroduced a direct indication about its future inclinations, suggesting that the systematic error may be strongly related to the communication strategy of the central bank.

In order to investigate the business-cycle properties and the predictability of the two different components $\theta_t^{(n)}$ and $\sigma_{t+n}^{(n)}$, we report in table 5 the results obtained from the following regressions for both the euro area and the United States:

$$\sigma_{t+n}^{(n)} = \alpha_\sigma^{(n)} + \beta_\sigma^{(n)} z_t + \gamma_\sigma^{(n)} f_t^{(n)} + \epsilon_{t+n}^{(n)} \tag{13}$$

$$\theta_t^{(n)} = \alpha_\theta^{(n)} + \beta_\theta^{(n)} z_t + \gamma_\theta^{(n)} f_t^{(n)} + \eta_t^{(n)}. \tag{14}$$

In both areas risk premia vary significantly along the business cycle. The coefficients of the business-cycle indicators are negative at all horizons and highly significant, and their magnitude increases with the forecast horizon. In periods of faster growth, risk premia in

[16] We use the same premium at all forecast horizons, assuming that $E_t[PR_{t+n}] = PR_t$ for $n = 1, \ldots, 6$.

[17] In order to investigate whether risk premia are, on average, different in the two areas, we run a regression, pooling the data of the two areas, on a constant and a country dummy. The estimated coefficients for the country dummy are not significantly different from zero at the 10 percent confidence level.

Table 5. Time-Varying Risk Premia

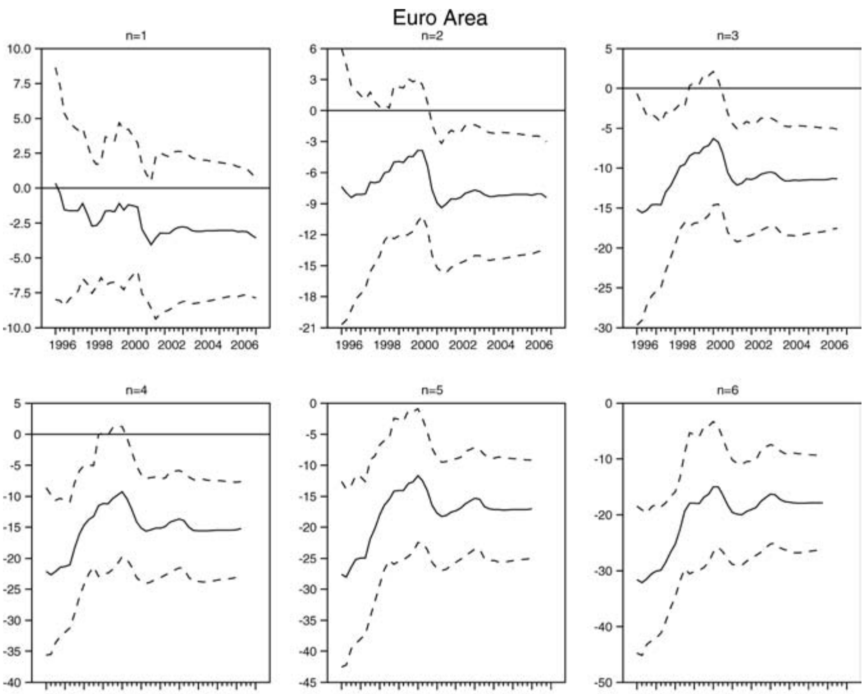
	$n=1$		$n=2$		$n=3$		$n=4$		$n=5$		$n=6$	
Euro Area												
Constant	9.1**	9.1**	13.6**	13.6**	17.7**	17.7**	24.5**	24.5**	30.4**	30.4**	34.2**	34.2**
	(1.9)	(1.9)	(3.3)	(2.9)	(4.3)	(3.9)	(5.4)	(5.0)	(5.6)	(5.1)	(4.3)	(5.9)
RGDP	−2.0	—	−4.6	—	−8.0*	—	−12.4**	—	−14.5**	—	−13.7**	—
	(1.9)		(3.8)		(4.4)		(5.4)		(5.8)		(5.8)	
E(empl)	—	−3.6*	—	−8.4**	—	−11.4**	—	−15.2**	—	−17.0**	—	−17.8**
		(2.0)		(3.5)		(4.0)		(4.5)		(4.9)		(4.9)
Future	7.7**	7.6**	16.0**	15.6**	23.5**	21.8**	32.3**	28.8**	38.6**	34.2**	41.0**	37.1**
	(2.3)	(2.0)	(4.3)	(3.3)	(5.1)	(3.6)	(6.1)	(4.1)	(6.1)	(3.9)	(6.6)	(4.4)
R^2	0.14	0.18	0.31	0.40	0.42	0.51	0.48	0.56	0.56	0.63	0.57	0.64
United States												
Constant	12.2**	12.2**	17.6**	17.6**	25.1**	25.1**	32.2**	32.2**	37.9**	37.9**	42.0**	42.0**
	(3.9)	(3.9)	(5.4)	(5.4)	(6.3)	(6.0)	(6.3)	(5.8)	(6.7)	(5.9)	(6.6)	(5.7)
RGDP	−4.7	—	−13.1**	—	−20.8**	—	−23.5**	—	−28.0**	—	−26.0**	—
	(4.8)		(6.6)		(7.6)		(7.5)		(7.8)		(7.5)	
NFP	—	−10.0**	—	−27.7**	—	−44.7**	—	−49.8**	—	−56.0**	—	−52.4**
		(7.8)		(10.4)		(11.1)		(10.3)		(9.7)		(8.8)
Future	10.1**	13.5*	21.4**	34.6**	32.9**	55.7**	41.2**	66.6**	52.2**	78.8**	58.9**	82.8**
	(4.3)	(7.6)	(6.0)	(10.0)	(6.8)	(10.6)	(6.9)	(9.9)	(7.0)	(9.2)	(6.7)	(8.3)
R^2	0.28	0.26	0.24	0.27	0.31	0.39	0.43	0.53	0.52	0.64	0.62	0.73

Notes: The sample period is 1994:Q1–2007:Q1. Ex ante risk premia are measured in basis points. Predictive regressions are estimated by OLS. Newey-West standard errors are reported in parentheses. * denotes significance at the 10 percent confidence level; ** denotes significance at the 5 percent confidence level.

the euro area may range between 10 basis points (for the one-quarter horizon) and 40 points (for the six-quarter horizon); in periods of slower (or negative) growth, they are between 20 and 80 basis points. In the United States, risk premia range from 10 to 25 basis points in periods of faster growth and from 25 to 95 basis points in periods of slower (or negative) growth.

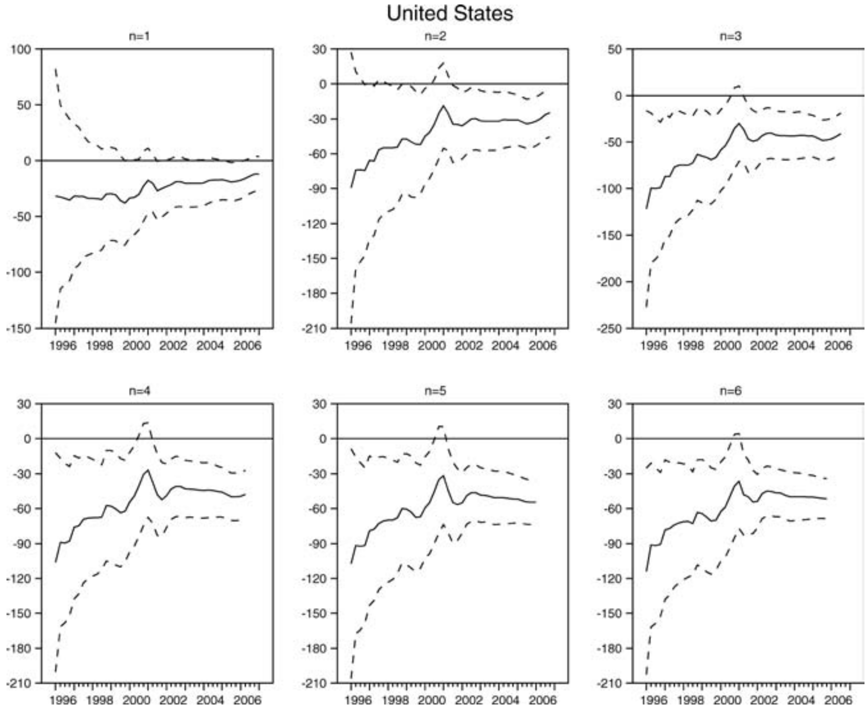
The recursive estimates of the risk-premia equation (figure 3) and the corresponding CUSUM tests (figure A4 in the appendix) suggest that the sign and the significance of the estimated relationships between risk premia and the business cycle (and, more in general, of the estimated regression) are stable over time in both areas. Moreover, as shown in table A1 in the appendix, unit-root tests suggest that risk premia are stationary at all horizons considered.

Figure 3. Recursive Coefficients for the Business-Cycle Indicator in Risk Premia Regressions



(continued)

Figure 3. (Continued)



Notes: Recursive least-squares estimates. The first estimate is obtained using the sample 1994:Q1–1996:Q1. Employment expectations for the months ahead and nonfarm payrolls are used respectively in predictive regressions for the euro area and for the United States. Dotted lines represent the two standard-error bands around the estimated coefficients.

As a robustness check for the euro area, we consider the shorter sample period 1999:Q1–2007:Q3 (table 6). The estimates suggest that with stage 3 of the EMU the risk premia have diminished in the euro area but have still remained statistically significant at all forecast horizons. Moreover, the coefficients of employment expectations are negative and highly significant at horizons beyond one quarter and they are of the same magnitude of those obtained in the overall sample.

The predictability of ex post prediction errors along the business cycle is assessed in table 7. The estimated relationships between forecast errors and business-cycle indicators largely resemble those

Table 6. Time-Varying Risk Premia in the Euro Area After the Start of Stage 3 of EMU

n	1	2	3	4	5	6
Constant	6.3**	6.5*	8.1*	10.8**	12.1**	12.4*
	(2.2)	(3.4)	(4.4)	(5.4)	(5.8)	(6.6)
E(empl)	−5.3	−9.4*	−14.0**	−18.9**	−19.7**	−17.0**
	(4.1)	(5.4)	(6.6)	(7.7)	(8.2)	(7.2)
Future	8.1**	12.4**	18.0**	23.3**	27.2**	27.7**
	(3.9)	(3.9)	(5.0)	(5.6)	(6.5)	(5.8)
R^2	0.08	0.13	0.23	0.33	0.39	0.40

Notes: The sample period is 1999:Q1–2007:Q3. Newey-West standard errors are reported in parentheses. * denotes significance at the 10 percent confidence level; ** denotes significance at the 5 percent level.

of total excess returns. In the euro area, employment expectations are not significantly correlated with forecast errors, while in the United States the estimated coefficients are significantly negative at all horizons.[18]

A theoretical analysis of the reasons behind the presence of forecast errors that are predictable and significantly countercyclical only in the United States lies beyond the scope of this paper. However, it should be noted that in the presence of structural changes, economic agents may need time to learn about the new environment: in the early stages of this process, previously held beliefs could lead to systematic biased predictions. To the extent that learning behaviors converge to rational expectations, the prediction bias would be a temporary phenomenon (see, for example, Evans and Honkapohja 2001). Therefore, it is not surprising that in the sample analyzed here the properties of the ex post prediction error are different in the two areas and change over time.

In this respect, a possible explanation for the empirical evidence described in this section regarding prediction errors in the United

[18]Bacchetta, Mertens, and van Wincoop (2006) analyze excess returns and forecast errors in the foreign exchange market and find that, in general, the predictability of the two measures are strictly related, in the sense that a variable that is successfully used in predicting expectation errors is also helpful for predicting the total excess returns.

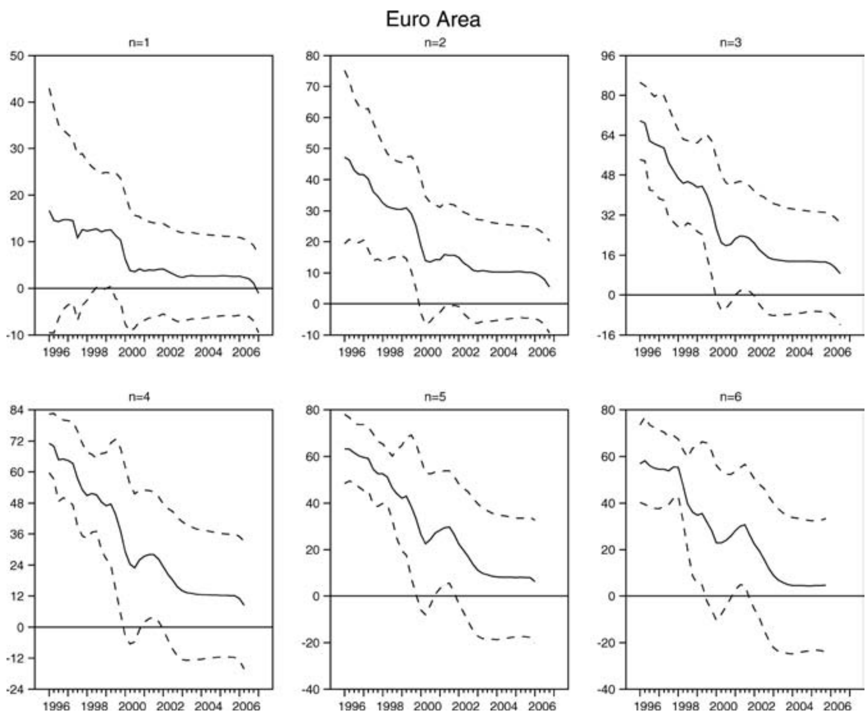
Table 7. Time-Varying Forecast Errors

	$n = 1$		$n = 2$		$n = 3$		$n = 4$		$n = 5$		$n = 6$	
Euro Area												
Constant	−0.7	−0.7	6.9	6.9	19.9	19.9	34.6*	34.6*	50.3**	50.3**	68.0**	68.0**
	(4.3)	(4.7)	(9.7)	(10.3)	(14.9)	(15.6)	(18.8)	(19.5)	(21.2)	(21.7)	(23.6)	(24.0)
RGDP	−11.5**	—	−13.5	—	−12.0	—	−10.1	—	−8.1	—	−7.1	—
	(4.4)		(12.6)		(19.6)		(21.6)		(20.5)		(19.1)	
E(empl)	—	−1.1	—	5.4	—	8.5	—	8.5	—	6.3	—	4.6
		(5.1)		(10.1)		(15.1)		(16.9)		(17.3)		(17.6)
Future	9.0*	2.7	16.5	7.3	28.7	20.0	46.7**	39.5**	62.4**	57.2**	74.9**	70.9**
	(4.9)	(3.7)	(13.5)	(8.8)	(19.0)	(12.7)	(20.2)	(14.5)	(19.9)	(16.1)	(20.4)	(18.0)
R^2	0.08	−0.03	0.04	0.00	0.07	0.07	0.17	0.17	0.28	0.28	0.34	0.33
United States												
Constant	6.1	6.1	15.7*	15.7*	26.6*	26.6*	41.5**	41.5**	55.7**	55.7**	70.2**	70.2**
	(5.2)	(5.2)	(9.5)	(9.2)	(14.0)	(13.4)	(17.3)	(15.8)	(20.1)	(17.6)	(22.0)	(18.9)
RGDP	−10.2	—	−25.0**	—	−37.5**	—	−53.5**	—	−62.7**	—	−74.0**	—
	(6.5)		(12.2)		(17.6)		(21.4)		(24.2)		(22.0)	
NFP	—	−16.9	—	−57.2**	—	−88.0**	—	−131.4**	—	−155.2**	—	−169.5**
		(11.3)		(20.4)		(28.1)		(31.1)		(31.6)		(31.2)
Future	1.9	13.9	13.9	52.5**	22.4	81.8**	41.1*	127.4**	55.4**	155.4**	69.6**	174.4**
	(6.3)	(11.3)	(12.1)	(20.4)	(17.4)	(28.1)	(21.2)	(31.1)	(23.9)	(31.4)	(25.0)	(30.7)
R^2	0.25	0.15	0.17	0.17	0.16	0.21	0.18	0.30	0.20	0.39	0.25	0.45

Notes: The sample period is 1994:Q1–2007:Q1. Ex post forecast errors are measured in basis points. Predictive regressions are estimated by OLS. Newey-West standard errors are reported in parentheses. * denotes significance at the 10 percent confidence level; ** denotes significance at the 5 percent confidence level.

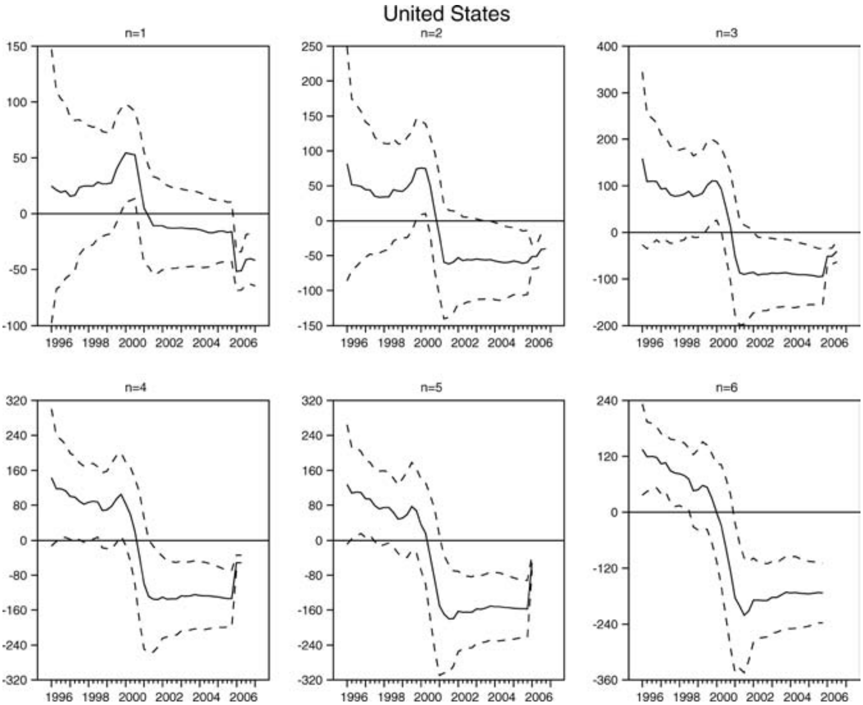
States may be the following. Throughout the decades of the 1990s, inflation and unemployment were trending down, while productivity was trending up. Forecasters and financial markets had a difficult time picking up on these developments in real time. As a result, they made repeated positive forecast errors in predicting inflation and negative ones in predicting output developments. As a consequence, forecast errors in predicting the future path of interest rates have been relatively small with respect to those realized in the 2000s (see figure A3 in the appendix), consistently with the assumption that the central bank sets interest rates in response to output and inflation (Taylor-type rules). Relatively small prediction errors

Figure 4. Recursive Coefficients for the Business-Cycle Indicator in the Ex Post Forecast Regressions



(continued)

Figure 4. (Continued)



Notes: Recursive least-squares estimates. The first estimate is obtained using the sample 1994:Q1–1996:Q1. Employment expectations for the months ahead are used in predictive regressions for the euro area. Nonfarm payrolls are used in predictive regressions for the United States. Dotted lines represent the two standard-error bands around the estimated coefficients.

were, therefore, associated with relatively high economic growth in the 1990s, coherently with our estimates.[19]

Figure 4 reports the recursive estimates of the coefficients of the business-cycle indicator used in equation (13) and shows that they have significantly decreased over time both in the euro area and in the United States, thus suggesting that the instability observed in the estimates of total excess return reflects the instability of the estimates of the ex-post systematic error (see also figure A5 in the appendix).

[19]We thank an anonymous referee for suggesting this point.

5. Out-of-Sample Forecast Accuracy

Insofar as risk premia and forecast errors are predictable by means of business-cycle indicators, it is interesting to investigate whether gains are achieved in out-of-sample forecast accuracy for short-term interest rates by using adjusted futures rates.

The design of the experiment is based on rolling-endpoint regressions. An initial estimate of risk premia at different horizons is obtained using the sample period 1994:Q1–1996:Q4; we use the estimate to compute a set of out-of-sample forecasts for future interest rates up to six quarters, as follows:

$$i^{f}_{t+n} = f^{(n)}_{t} - E_t\left(\widehat{x}^{(n)}_{t+n}\right). \tag{15}$$

We then add a new observation and repeat the forecasting exercise, until the end of the sample period. Overall we collect a set of fifty-eight out-of-sample predictions at each forecast horizon. In table 8 we report the mean error (ME) and the root-mean-squared errors (RMSEs) for (i) futures rates adjusted for time-varying risk premia, (ii) constant-adjusted futures rates, and (iii) futures rates adjusted for time-varying total excess return. We perform a Diebold-Mariano test to check whether the errors obtained under the adjusted predictions are significantly different from their counterparts obtained with unadjusted futures rates.

Unadjusted futures rates perform relatively poorly in both areas. In the euro area the RMSEs of the predictions obtained with the unadjusted futures rates are larger than those obtained from a random-walk model at all horizons beyond three quarters and those obtained from the Consensus Forecast survey at all horizons beyond one quarter. Futures rates adjusted for a constant excess return already produce lower RMSEs at all forecast horizons, even if the gains in forecast accuracy are small and often not significant (RMSE is reduced by about 10 to 25 percent with respect to that obtained with unadjusted futures). Adjusting futures rates for the time-varying risk premia further improves our predictions (by about 10 percent compared with those obtained with constant-unadjusted futures). Finally, adjusting for the time-varying excess return reduces the RMSE with respect to that obtained adjusting only for the risk premia by about 5 to 25 percent at horizons longer

Table 8. Out-of-Sample Forecasts for Short-Term Interest Rates: Summary Statistics

	$n = 1$		$n = 2$		$n = 3$		$n = 4$		$n = 5$		$n = 6$	
	ME	RMSE	ME	RMSE	ME	RMSE	ME	RMSE	ME	RMSE	ME	RMSE
Euro Area												
Random Walk	−3.6	32.5	−7.8	54.5	−13.6	73.1	−19.4	90.0	−24.8	102.6	−30.1	112.4
Consensus	1.6	29.7	5.8	49.5	19.2	72.2	36.0	93.3	53.3	110.8	73.2	128.4
Unadjusted	9.6	25.9	23.4	51.4	42.3	80.6	66.3	113.4	92.5	141.5	118.2	165.5
Constant Adj.	−10.0	23.4	−20.1	45.4	−30.7	68.9	−43.2	94.1	−52.6	112.0	−59.2	122.4
Risk-Premia Adj.	3.2	21.6	9.2	40.5	22.9	61.3	40.6	83.3	53.6	99.0	70.6	113.4
Excess Returns Adj.	−2.1	22.3	−4.8	44.3	−5.9	65.3	2.1	79.1	11.5	85.4	20.7	89.1
United States												
Random Walk	1.5	49.5	2.4	85.4	0.2	116.4	−2.9	146.0	−6.2	172.6	−11.0	194.7
Consensus	9.5	36.7	19.6	68.8	30.9	101.5	47.3	131.2	63.5	156.2	79.8	176.6
Unadjusted	11.2	39.7	27.3	77.1	46.0	115.9	68.9	154.3	91.0	188.7	112.5	217.3
Constant Adj.	1.7	36.9	1.5	72.6	3.4	109.9	2.0	146.4	3.4	178.3	9.4	203.3
Risk-Premia Adj.	1.7	37.4	12.4	72.7	28.1	103.1	48.0	135.0	69.4	161.3	90.6	182.1
Excess Returns Adj.	−0.4	41.7	2.1	71.7	12.2	98.1	16.9	113.0	19.9	120.4	23.2	128.9

Notes: ME is the mean error; RMSE is the root-mean-squared error. Forecast errors are measured in basis points. Employment expectations for the months ahead are used in predictive regressions for the euro area. Nonfarm payrolls are used in predictive regressions for the United States.

than three quarters; however, at shorter horizons there are no significant improvements, thus confirming that in the sample analyzed here the forecast errors are not predictable by means of business-cycle indicators and are, on average, not significant at shorter horizons.

For the United States, adjusting for the time-varying excess returns improves our forecasts by up to 40 percent with respect to unadjusted futures rates, while futures rates adjusted only for the risk premia determine RMSEs between 10 and 40 percent larger than those obtained adjusting for the total excess return at horizons longer than one quarter. In this case, prediction errors are significant and predictable by means of business-cycle indicators.

These results have important implications for central banks. Even if futures rates adjusted for both risk premia and systematic prediction errors are the best predictors of future monetary policy decisions at least at longer horizons, they no longer coincide with financial markets' expectations. Therefore, for a correct assessment of the financial markets' view about future policy decisions, policymakers should use quoted futures rates adjusted only for risk premia, as systematic forecast errors represent part of agents' expectations formation process.

6. Conclusions

In this paper we show that the prices of futures contracts on three-month interest rates are biased forecasts of future short-term interest rates. We also find evidence of large and time-varying excess returns on three-month interest rate futures in the euro area, in line with the results obtained by Piazzesi and Swanson (2004) for the United States. However, unlike those in dollars, ex post excess returns on futures contracts in euros do not appear to be significantly related to business-cycle indicators, while in both areas the sign and the significance of the estimated relationships between excess returns and the business cycle are unstable over time.

We show that ex post excess returns can be divided into two components. The first is the effective ex ante risk premium demanded by investors when they buy or sell the financial contract. The second is an ex post systematic forecast error.

The empirical analysis reveals that the risk premia on futures contracts in euros and in dollars are not significantly different and, interestingly, they are significantly countercyclical in both areas. Moreover, the sign and the significance of the estimated relationships between risk premia and the business cycle turn out to be stable over time.

Finally, we find that the instability observed in the estimates of total excess returns in both areas and the lack of a significative relationship between that variable and business-cycle indicators in the euro area are determined by the instability of the estimates of the ex post systematic-error component.

The policy implication of our findings is that even though futures rates adjusted for both components are better forecasts of future monetary policy actions, in assessing markets' view about future policy decisions, it is better to use futures rates adjusted only by risk premia, as systematic forecast errors are part of agents' expectations.

Appendix. Tables and Figures

Table A1. Unit-Root Test for Risk Premia

	Euro Area		United States	
	No Trend	**Linear Trend**	**No Trend**	**Linear Trend**
$\theta_t^{(1)}$	−6.449**	−7.595**	−3.139**	−4.339**
$\theta_t^{(2)}$	−3.791**	−4.948**	−2.924**	−4.297**
$\theta_t^{(3)}$	−2.729**	−4.267**	−2.422**	−3.980**
$\theta_t^{(4)}$	−2.765**	−4.063**	−2.394**	−3.884**
$\theta_t^{(5)}$	−2.586**	−4.286**	−2.331**	−3.763**
$\theta_t^{(6)}$	−2.228**	−3.973**	−2.435**	−4.109**

Notes: The sample period is 1994:Q1–2007:Q1. The t-statistic of the augmented Dickey-Fuller (ADF) test includes in the test regression deterministic variables and p lagged difference terms of the dependent variable. The lag order p has been selected using a Schwarz information criterion with the maximum lag length of 8. * denotes the rejection of the null hypothesis at the 10 percent level; ** denotes the rejection of the null hypothesis at the 5 percent confidence level.

Figure A1. CUSUM Test of Instability for Excess Returns Regressions

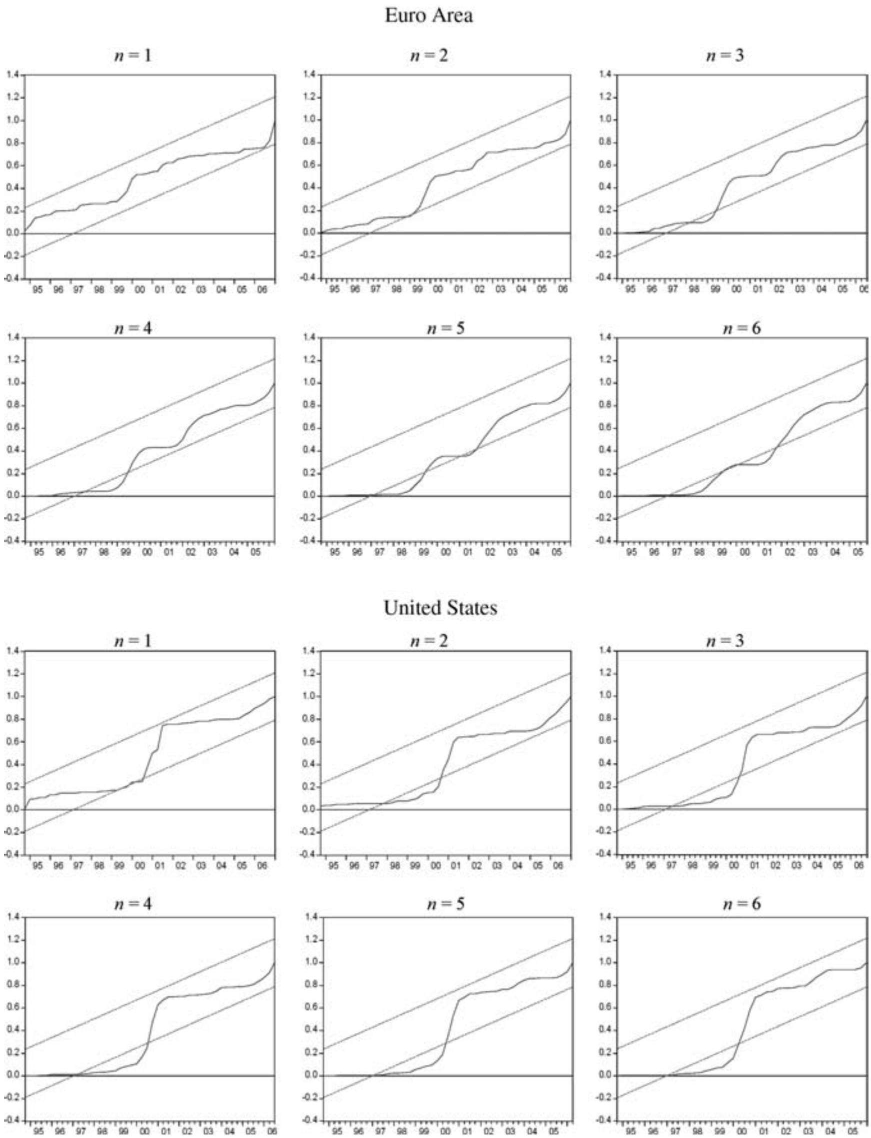


Figure A2. Risk Premia and Forecast Errors in the Euro Area

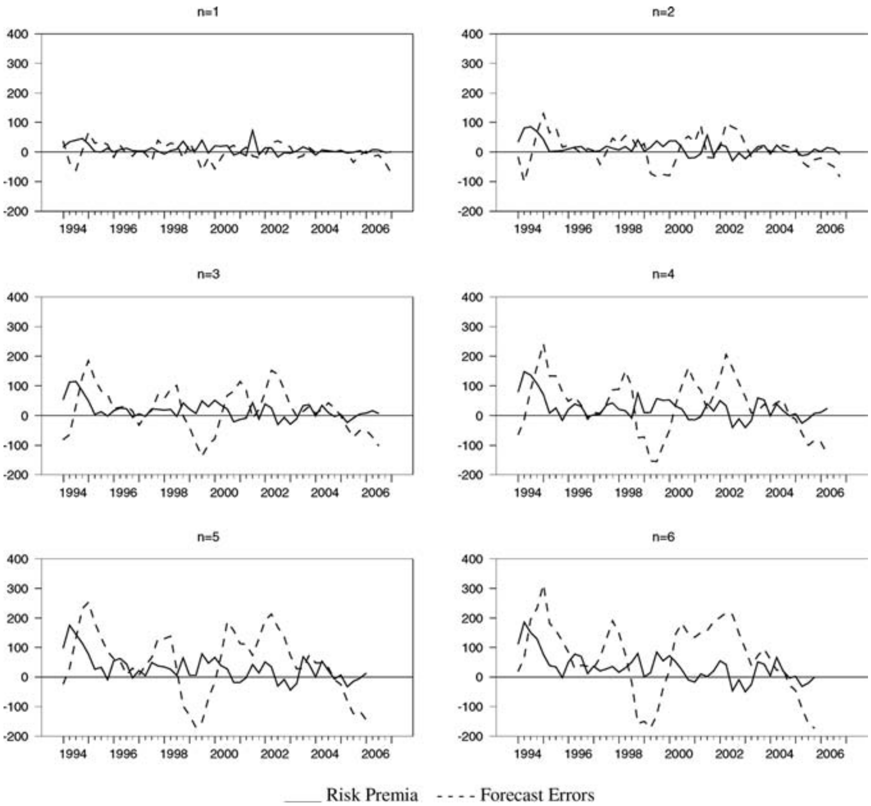


Figure A3. Risk Premia and Forecast Errors in the United States

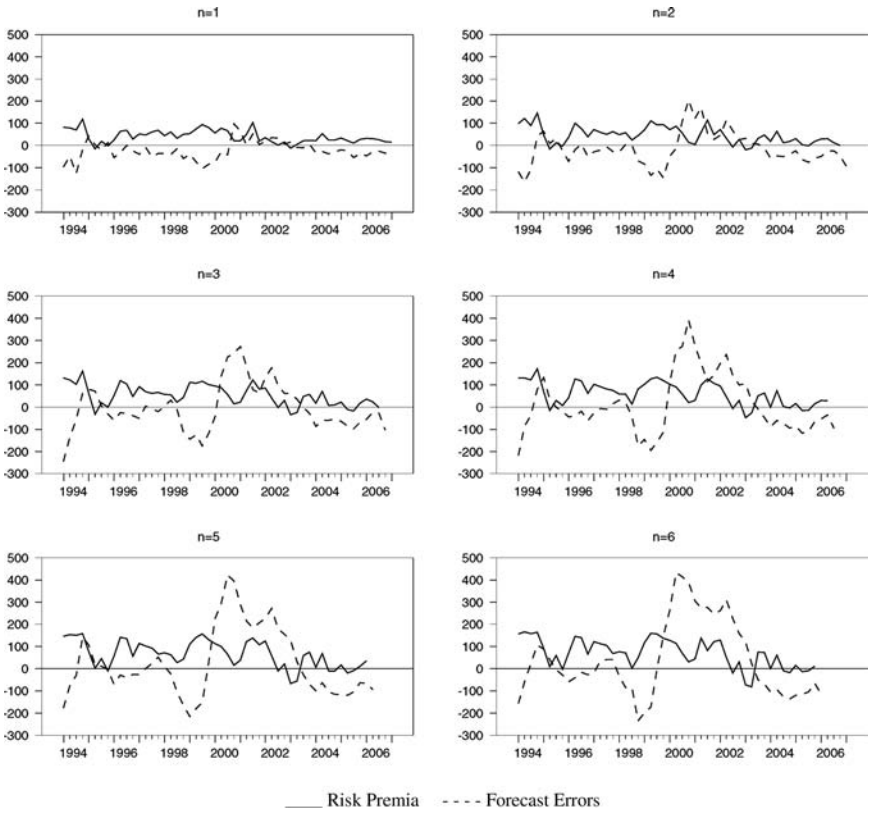


Figure A4. CUSUM Test of Instability for Risk Premia Regressions

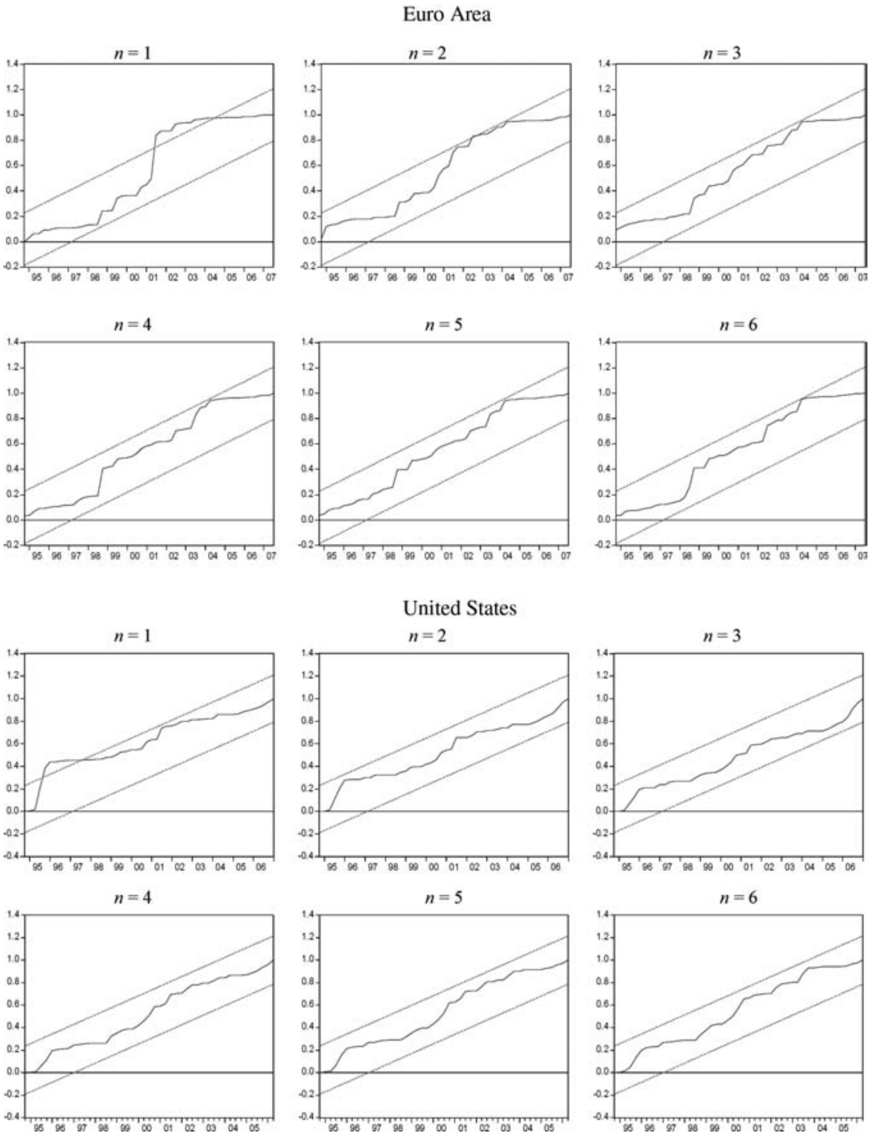
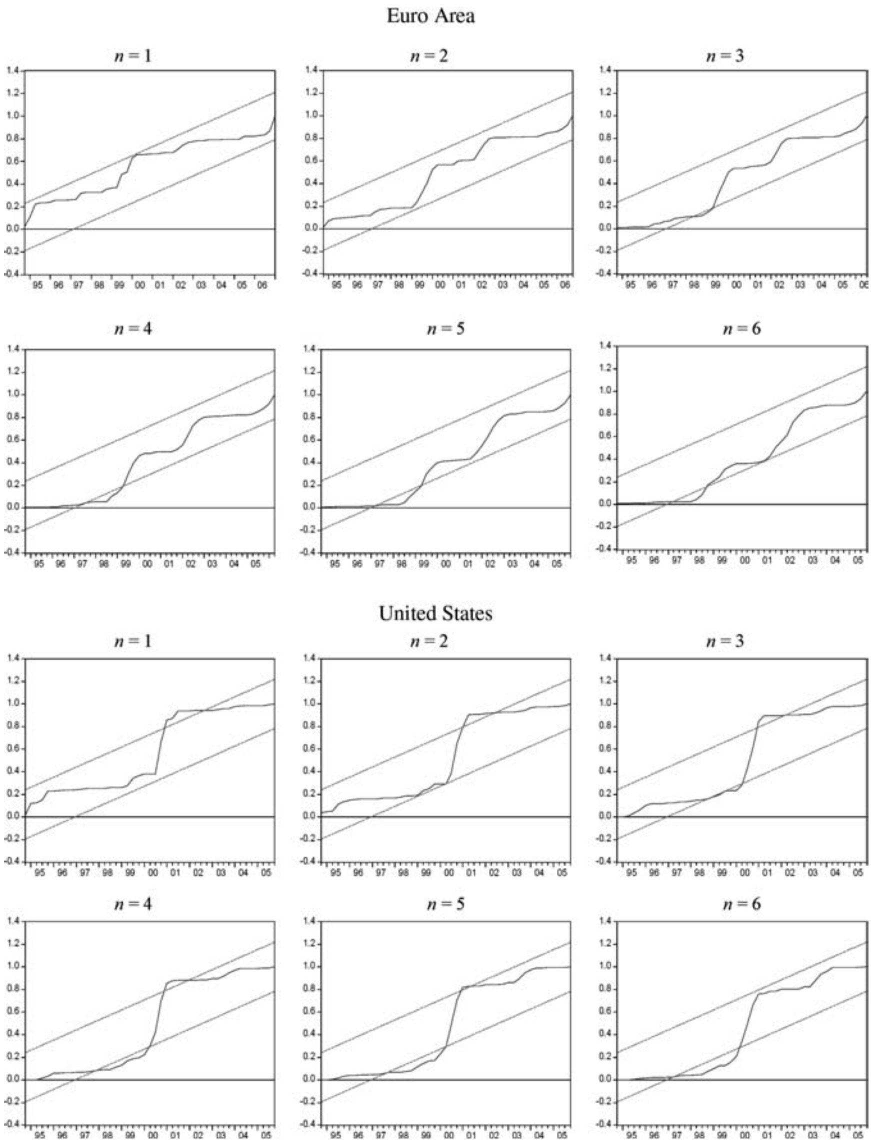


Figure A5. CUSUM Test of Instability for Forecast Errors Regressions



References

Ang, A., S. Dong, and M. Piazzesi. 2004. "No-Arbitrage Taylor Rules." Mimeo, Columbia University.

Ang, A., and M. Piazzesi. 2003. "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables." *Journal of Monetary Economics* 50 (4): 745–87.

Ang, A., M. Piazzesi, and M. Wei. 2006. "What Does the Yield Curve Tell Us about GDP Growth?" *Journal of Econometrics* 131 (1–2): 359–403.

Bacchetta, P., E. Mertens, and E. van Wincoop. 2006. "Predictability in Financial Markets: What Do Survey Expectations Tell Us?" CEPR Discussion Paper No. 5770.

Bekaert, G., R. J. Hodrick, and D. A. Marshall. 2001. "Peso Problem Explanations for Term Structure Anomalies." *Journal of Monetary Economics* 48 (2): 241–70.

Buiter, W. H. 1999. "Six Months in the Life of the Euro: What Have We Learnt?" *Tijdschrift voor Politieke Economie* 21 (4): 4–25.

Campbell, J. Y. 1995. "Some Lessons from the Yield Curve." *Journal of Economic Perspectives* 9 (3): 129–52.

Campbell, J. Y., and R. J. Shiller. 1991. "Yield Spreads and Interest Rate Movements: A Bird's Eye View." *Review of Economic Studies* 58 (3): 495–514.

Cochrane, J. H. 2006. "Financial Markets and the Real Economy." Graduate School of Business, University of Chicago.

Cochrane, J. H., and M. Piazzesi. 2002. "The Fed and Interest Rates: A High-Frequency Identification." *American Economic Review* 92 (2): 90–95.

Diebold, F. X., M. Piazzesi, and G. D. Rudebusch. 2005. "Modeling Bond Yields in Finance and Macroeconomics." *American Economic Review, Papers and Proceedings* 95 (2).

Durham, J. B. 2003. "Estimates of the Term Premium on Near-Dated Federal Funds Futures Contracts." FEDS Working Paper No. 2003-19, Board of Governors of the Federal Reserve System.

Elliott, G., T. J. Rothenberg, and J. H. Stock. 1996. "Efficient Tests for an Autoregressive Unit Root." *Econometrica* 64 (4): 813–36.

Evans, G., and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics.* Princeton, NJ: Princeton University Press.

Fama, E. F. 1984. "The Information in the Term Structure." *Journal of Financial Economics* 13 (4): 509–28.

———. 1991. "Efficient Capital Markets: II." *Journal of Finance* 46 (5): 1575–1617.

Fama, E. F., and R. R. Bliss. 1987. "The Information in Long-Maturity Forward Rates." *American Economic Review* 77 (4): 680–92.

Froot, K. A. 1989. "New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates." *Journal of Finance* 44 (2): 283–305.

Gourinchas, P., and A. Tornell. 2004. "Exchange Rate Puzzles and Distorted Beliefs." *Journal of International Economics* 64: 303–33.

Gürkaynak, R. S., B. Sack, and E. Swanson. 2006. "Market-Based Measures of Monetary Policy Expectations." Federal Reserve Bank of San Francisco Working Paper No. 2006-04.

Hördal, P., O. Tristani, and D. Vestin. 2006. "A Joint Econometric Model of Macroeconomic and Term-Structure Dynamics." *Journal of Econometrics* 131 (1–2): 405–44.

Johansen, S. 1991. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica* 59 (6): 1551–80.

———. 1995. "A Statistical Analysis of Cointegration for I(2) Variables." *Econometric Theory* 11 (1): 25–59.

Jongen, R., W. F. C. Verschoor, and C. C. Wolff. 2005. "Time Variation in Term Premia: International Evidence." CEPR Discussion Paper No. 4959.

Kim, D. H., and A. Orphanides. 2005. "Term Structure Estimation with Survey Data on Interest Rate Forecasts." FEDS Working Paper No. 2005-48, Board of Governors of the Federal Reserve System.

Krueger, J. T., and K. N. Kuttner. 1996. "The Fed Funds Futures Rate as a Predictor of Federal Reserve Policy." *Journal of Futures Markets* 16 (8): 865–79.

Pericoli, M., and M. Taboga. 2006. "Canonical Term-Structure Models with Observable Factors and the Dynamics of Bond Risk Premiums." Banca d'Italia Temi di Discussione No. 580.

Phillips, P. C. B., and B. E. Hansen. 1990. "Statistical Inference in Instrumental Variables Regression with I(1) Processes." *Review of Economic Studies* 57 (1): 99–125.

Piazzesi, M., and E. Swanson. 2004. "Futures Prices as Risk-Adjusted Forecasts of Monetary Policy." NBER Working Paper No. 10547.

Rapach, D. E., and C. E. Weber. 2004. "Are Real Interest Rates Really Nonstationary? New Evidence from Tests with Good Size and Power." *Journal of Macroeconomics* 26 (3): 409–30.

Rose, A. K. 1988. "Is the Real Interest Rate Stable?" *Journal of Finance* 43 (5): 1095–1112.

Rudebusch, G., and T. Wu. 2004. "A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy." *Proceedings* (Federal Reserve Bank of San Francisco) (March).

Sack, B. 2002. "Extracting the Expected Path of Monetary Policy from Futures Rates." FEDS Working Paper No. 2002-56, Board of Governors of the Federal Reserve System.

Sargent, T. J. 1986. *Rational Expectations and Inflation.* New York: Harper and Row.

Shiller, R. J. 2000. *Irrational Exuberance.* Princeton, NJ: Princeton University Press.

Sims, C. A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50 (3): 665–90.

Stock, J. H., and M. W. Watson. 1993. "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems." *Econometrica* 61 (4): 783–820.

Timmermann, A. 1993. "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices." *Quarterly Journal of Economics* 108 (4): 1135–45.

Firm-Specific Capital and Welfare*

Tommy Sveen[a] and Lutz Weinke[b,c]
[a]Monetary Policy Department, Norges Bank
[b]Department of Economics, Duke University
[c]Institute for Advanced Studies, Vienna

What are the consequences for monetary policy design implied by the fact that price setting and investment typically take place simultaneously at the firm level? To address this question we analyze simple (constrained) optimal interest rate rules in the context of a dynamic New Keynesian model featuring firm-specific capital accumulation as well as sticky prices and wages à la Calvo. We make the case for Taylor-type rules. They are remarkably robust in the sense that their welfare implications do appear to hinge neither on the specific assumptions regarding capital accumulation that are used in their derivation nor on the particular definition of natural output that is used to construct the output gap.

JEL Codes: E22, E31, E52.

1. Introduction

How does *firm-specific* capital accumulation affect the desirability of alternative arrangements for the conduct of monetary policy? We address this question employing a New Keynesian (NK) framework, i.e., a dynamic stochastic general equilibrium model featuring nominal rigidities combined with monopolistic competition. Specifically,

*Thanks to seminar participants at the 2006 Annual Congress of the EEA, Duke University, HEC Montréal, Humboldt-Universität zu Berlin, Ludwig-Maximilians-Universität München, Magyar Nemzeti Bank, Norges Bank, Universität Bonn, Université du Luxembourg, Universität Trier, Universität zu Köln, and Universität Würzburg. Special thanks to Ida Wolden Bache, Jordi Galí, Øistein Røisland, and Stephanie Schmitt-Grohé, as well as to the editors Carl Walsh and Michael Woodford and two anonymous referees. The usual disclaimer applies. The views expressed in this paper are those of the authors and should not be attributed to Norges Bank. E-mails: tommy.sveen@norges-bank.no and weinkel@econ.duke.edu.

we consider an economic environment with sticky prices and wages à la Calvo (1983). Our model is therefore similar to the one developed in Erceg, Henderson, and Levin (2000) except for the fact that we allow for capital accumulation.[1] The welfare criterion is derived from the average of household utility functions, along the lines of Rotemberg and Woodford (1998).

What is the relevance of our analysis? Edge (2003) shows how the work by Rotemberg and Woodford (1998) can be extended to conduct a welfare analysis in the context of an NK model where capital accumulation is endogenous. She assumes, however, that firms have access to a rental market for capital,[2] which is not an innocuous simplification in an NK model, as analyzed in Sveen and Weinke (2005, 2007) and Woodford (2005).[3] In the present paper we show how a welfare analysis can be conducted in the context of an NK model featuring firm-specific capital accumulation (FS for short). Moreover, we explain how and why the conclusions regarding the desirability of monetary policy change if a rental market for capital (RM for short) is assumed instead. The latter analysis is of considerable interest given the widespread use of the rental-market assumption in the context of welfare-based evaluations of monetary policy.[4]

We obtain three results. First, we analyze interest rate rules according to which the central bank adjusts the nominal interest rate only in response to changes in price and wage inflation. In that context we find that optimized interest rate rules prescribe putting relatively more weight on price inflation than on wage inflation under FS, whereas the opposite is true under RM. Moreover, using the optimized interest rate rule associated with RM in the FS specification implies a large welfare loss, as we discuss. Interestingly, the

[1] Erceg, Henderson, and Levin (2000) assume that the aggregate capital stock is constant and that there exists a rental market for capital.

[2] Another difference between our work and Edge's is that she assumes frictionless investment, whereas we follow Woodford (2005) in assuming a convex adjustment cost at the firm level.

[3] Schmitt-Grohé and Uribe (2007b) argue that both the rental-market assumption and the assumption of firm-specific capital are somewhat extreme. However, the work by Altig et al. (2005) and Eichenbaum and Fisher (2007) suggests that the assumption of firm-specific capital is appealing on empirical grounds.

[4] See, e.g., Levin et al. (2006) and Schmitt-Grohé and Uribe (2007b).

additional endogenous price stickiness implied by the presence of firm-specific capital (and the lack thereof under RM) is identified as the main reason behind the difference in implications for optimal monetary policy design. Sveen and Weinke (2005) show that this is the *only* difference between the two models if attention is restricted to a first-order approximation to equilibrium dynamics. In the present paper we therefore find that our price stickiness metric is also useful from a normative point of view. This is surprising because our welfare criterion, a second-order approximation to the unconditional expectation of the household's utility, is *not* identical in the two models if the price stickiness is increased in RM in such a way that FS and RM are identical, up to the first order. Let us relate our first result to the existing literature. Schmitt-Grohé and Uribe (2007a) demonstrate in the context of a rental-market model that the relative weight attached to price and wage inflation in an optimized interest rate rule depends crucially on which nominal variable is stickier.[5] Our analysis shows that the difference in policy implications between FS and RM can be understood in an analogous way.

We also analyze Taylor-type rules, i.e., interest rate rules prescribing that the central bank reacts to price inflation and to the output gap. Our second result is that those interest rate rules are remarkably robust in the following sense. If the optimized rule implied by RM is used in FS, then the resulting welfare loss is small compared with the outcome under the optimized rule associated with that model.[6] Consequently, the central bank does not need to take a stand on which specification of capital accumulation is the empirically more plausible one if it uses a Taylor-type rule.

But how should the output gap be defined? So far there is no consensus in the literature on the answer to that question. Neiss

[5]Schmitt-Grohé and Uribe (2006) make the case for price stability as the central goal of optimal monetary policy. They show that desirable outcomes can be implemented by a combination of passive monetary and active fiscal policy. In the present paper we focus exclusively on optimal monetary policy.

[6]In related work Levin, Wieland, and Williams (1999) and Levin and Williams (2003) argue that simple Taylor-type rules optimized for one model perform well under alternative models. Our work extends their analysis to a setting that takes into account how the microfounded welfare criterion depends on the model that is used.

and Nelson (2003) and Woodford (2003, ch. 5) propose two alternative definitions. Our third result is that the difference between these two competing definitions matters very little for the resulting welfare implications, and we explain why this is so.

The remainder of the paper is organized as follows. The model is outlined in section 2. We present the welfare criterion in section 3. Our results are shown and interpreted in section 4. Section 5 concludes.

2. The Model

The model we use to analyze the implications of firm-specific capital accumulation for monetary policy design is an NK framework with complete financial markets.

2.1 Households

Households maximize expected discounted utility

$$E_t \sum_{k=0}^{\infty} \beta^k U(C_{t+k}, N_{t+k}(h)),$$

where β is the subjective discount factor. The subscript t is generally used to indicate that a variable is dated as of that period and E_t denotes an expectation operator that is conditional on information available through time t. Hours worked by household h are given by $N_t(h)$ and C_t is a Dixit-Stiglitz consumption aggregate. Specifically,

$$C_t \equiv \left(\int_0^1 C_t(i)^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}, \tag{1}$$

where ε is the elasticity of substitution between different varieties of goods $C_t(i)$. The associated price index is defined as $P_t \equiv (\int_0^1 P_t(i)^{1-\varepsilon} di)^{\frac{1}{1-\varepsilon}}$ and optimal allocation of any spending on the available goods implies that consumption expenditure can be written as $P_t C_t$. Household h's period utility is given by the following function:

$$U(C_t, N_t(h)) = \frac{C_t^{1-\sigma}}{1-\sigma} - \frac{N_t(h)^{1+\phi}}{1+\phi}, \tag{2}$$

where parameter σ denotes the household's relative risk aversion and parameter ϕ can be interpreted as the inverse of the Frisch aggregate labor-supply elasticity. Each household is assumed to be the monopolistically competitive supplier of its differentiated type of labor, $N_t(h)$. As in Erceg, Henderson, and Levin (2000), we assume staggered wage setting à la Calvo (1983); i.e., each household faces a constant and exogenous probability, θ_w, of getting to reoptimize its wage in any given period. Our assumptions of separable preferences combined with complete financial markets imply that the heterogeneity across households in their hours worked does not translate into consumption heterogeneity. This is reflected in our notation. Optimizing behavior on the part of firms implies that demand for type h labor, $N_t^d(h)$, is given by

$$N_t^d(h) = \left(\frac{W_t(h)}{W_t} \right)^{-\varepsilon_N} N_t^d, \tag{3}$$

where $W_t(h)$ denotes the nominal wage posted by household h and ε_N gives the elasticity of substitution between different types of labor. Finally, W_t and N_t^d denote, respectively, the aggregate nominal wage and aggregate labor demand. They are defined as the corresponding aggregate prices and quantities for goods.

Under standard assumptions, the relevant budget constraint prescribes that the present value of all expenditures cannot be greater than the value of a household's initial assets and the present value of its income. The latter derives from wage payments and profits resulting from ownership of firms net of taxes.[7] We assume that there are only lump-sum taxes and the only role of the government is to levy these taxes to finance subsidies in goods and factor markets which render the steady state of our model Pareto optimal. This assumption in turn is needed to compute our welfare criterion up to the second order using a first-order approximation to the equilibrium dynamics, as we are going to see.

For future reference, let us note two implications of households' optimizing behavior. First, we obtain a stochastic discount factor for random nominal payments, $Q_{t,t+1}$, from a standard intertemporal optimality condition,

[7]For details, see Woodford (2003, ch. 2).

$$\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \left(\frac{P_t}{P_{t+1}} \right) = Q_{t,t+1}. \tag{4}$$

The stochastic discount factor is linked to the gross nominal interest rate, R_t, by the relationship $E_t\{Q_{t,t+1}\} = R_t^{-1}$, which holds in equilibrium. Second, under our assumptions the first-order condition for wage setting reads

$$E_t^w \left\{ \sum_{k=0}^{\infty} (\beta\theta_w)^k \, N_{t+k}^d(h) \, C_{t+k}^{\sigma} \left[\frac{W_t(h)}{P_{t+k}} - MRS_{t+k}(h) \right] \right\} = 0, \tag{5}$$

where $MRS_t(h) \equiv N_t(h)^{\phi} \; C_t^{\sigma}$ is the marginal rate of substitution of consumption for leisure of household h. Moreover, E_t^w is meant to indicate an expectation that is conditional on time t information, but integrating only over those future states in which the household has not reset its wage since period t.

2.2 Firms

There is a continuum of firms, and each of them is the monopolistically competitive producer of a differentiated good. Each firm i has access to a Cobb-Douglas technology,

$$Y_t(i) = X_t K_t(i)^{\alpha} N_t(i)^{1-\alpha}, \tag{6}$$

where parameter α measures the capital share in the production function. Aggregate technology is given by X_t, and $K_t(i)$ and $N_t(i)$ denote, respectively, firm i's capital stock and labor input used in its production $Y_t(i)$. Technology shocks are assumed to be the only source of aggregate uncertainty.[8] This is another modeling choice that is guided by Erceg, Henderson, and Levin (2000).[9] Specifically, we consider a stationary $AR(1)$ process for the log of technology,

$$x_t = \rho_a x_{t-1} + \varepsilon_t, \tag{7}$$

[8] Of course, the extent to which technology shocks are an important source behind the observable business-cycle fluctuations is the topic of an ongoing debate. See, e.g., Galí and Rabanal (2005).

[9] Strictly speaking, Erceg, Henderson, and Levin (2000) do not only assume technology shocks, but they restrict their welfare analysis to this kind of shock.

where parameter $\rho_a \in (0,1)$ and ε_t is assumed to be i.i.d. with mean zero. Firms face three additional restrictions. First, we assume Calvo (1983) pricing; i.e., each firm faces a constant and exogenous probability, θ, of getting to reoptimize its price in any given period. Second, we follow Woodford (2005) in assuming that investment at the firm level is restricted in the following way:

$$I_t(i) = \Gamma\left(\frac{K_{t+1}(i)}{K_t(i)}\right) K_t(i). \tag{8}$$

In the last equation, $I_t(i)$ denotes the amount of the composite good[10] necessary to change firm i's capital stock from $K_t(i)$ to $K_{t+1}(i)$ one period later. Moreover, function $\Gamma(\cdot)$ is assumed to satisfy the following: $\Gamma(1) = \delta$, $\Gamma'(1) = 1$, and $\Gamma''(1) = \epsilon_\psi$. Parameter δ denotes the depreciation rate and $\epsilon_\psi > 0$ measures the convex capital adjustment cost in a log-linear approximation to the equilibrium dynamics. Third, cost minimization by firms and households implies that demand for each individual good i can be written as follows:

$$Y_t^d(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\varepsilon} Y_t^d, \tag{9}$$

where $Y_t^d \equiv C_t + I_t$ denotes aggregate demand and $I_t \equiv \int_0^1 I_t(i)di$ denotes aggregate investment demand. Given those constraints, each firm i is assumed to maximize its market value:

$$\max \sum_{k=0}^{\infty} E_t\left\{Q_{t,t+k}\left[Y_{t+k}^d(i)P_{t+k}(i) - W_{t+k}N_{t+k}(i) - P_{t+k}I_{t+k}(i)\right]\right\}.$$

For future reference, let us mention two implications of optimizing behavior at the firm level. First, firm i's first-order condition for capital accumulation reads

$$\Gamma_t'(\cdot)P_t = E_t\left\{Q_{t,t+1}P_{t+1}\left[MS_{t+1}(i) - \Gamma_{t+1}(\cdot) + \Gamma_{t+1}'(\cdot)\frac{K_{t+2}(i)}{K_{t+1}(i)}\right]\right\}, \tag{10}$$

[10] We assume that the elasticity of substitution is the same as in the consumption aggregate.

where $MS_t(i) \equiv \frac{W_t}{P_t}\frac{MPK_t(i)}{MPL_t(i)}$ is the real marginal return to firm i's capital. The latter results from savings in labor costs. Second, let us note that under our assumptions firm i's first-order condition for price setting reads

$$\sum_{k=0}^{\infty} \theta^k E_t^p \left\{ Q_{t,t+k} Y_{t+k}^d(i) \left[P_t^*(i) - P_{t+k} MC_{t+k}^n(i) \right] \right\} = 0, \tag{11}$$

where $MC_t(i) \equiv \frac{W_t/P_t}{MPL_t(i)}$ measures the real marginal cost, $MPL_t(i)$ is the marginal product of labor of firm i, and E_t^p is meant to indicate an expectation that is conditional on time t information, but integrating only over those future states in which the firm has not reoptimized its price since period t.

2.3 Market Clearing

Clearing of the labor market requires for each type of labor h

$$N_t(h) = N_t^d(h). \tag{12}$$

Likewise, market clearing for each variety i requires at each point in time

$$Y_t(i) = C_t^d(i) + I_t^d(i), \tag{13}$$

where $C_t^d(i)$ is consumption demand for good i, while $I_t^d(i)$ denotes investment demand for that good. To close the model, we need to specify monetary policy. We will come back to that point.

2.4 Some Linearized Equilibrium Conditions

The starting point of our welfare analysis is a linear approximation to the equilibrium dynamics around a steady state with zero inflation. Since the details of the linearization have been developed elsewhere,[11] we just briefly mention the resulting equilibrium conditions. Unless specified otherwise, a lowercase letter denotes the log-deviation of the original variable from its steady-state value. Let us

[11] See, e.g., Erceg, Henderson, and Levin (2000), Sveen and Weinke (2005), and Woodford (2005).

already note that the linearized equilibrium conditions are identical for FS and RM, except for the respective inflation equations.

The consumption Euler equation reads

$$c_t = E_t c_{t+1} - \frac{1}{\sigma}(r_t - E_t \pi_{t+1} - \rho), \tag{14}$$

where $\rho \equiv -\log \beta$ is the time discount rate. We have also used the notation $r_t \equiv \log(R_t)$ for the nominal interest rate and $\pi_t \equiv \log(\frac{P_t}{P_{t-1}})$ for inflation. The law of motion of capital is obtained from averaging and aggregating optimizing investment decisions on the part of firms. This implies

$$\Delta k_{t+1} = \beta E_t \Delta k_{t+2} + \frac{1}{\epsilon_\psi}[(1-\beta(1-\delta))E_t ms_{t+1} - (r_t - E_t \pi_{t+1} - \rho)], \tag{15}$$

where $K_t \equiv \int_0^1 K_t(i)di$ denotes aggregate capital, and $MS_t \equiv \int_0^1 MS_t(i)di$ measures the average real marginal return to capital. Aggregate production is pinned down by aggregate labor, capital, and technology:

$$y_t = x_t + \alpha k_t + (1-\alpha)n_t. \tag{16}$$

The wage-inflation equation results from averaging and aggregating optimal wage-setting decisions on the part of households. It takes the following simple form:

$$\omega_t = \beta E_t \omega_{t+1} + \lambda_\omega (mrs_t - rw_t), \tag{17}$$

where $\lambda_\omega \equiv \frac{(1-\beta\theta_w)(1-\theta_w)}{\theta_w}\frac{1}{1+\eta\varepsilon_N}$. Moreover, $\omega_t \equiv \log(\frac{W_t}{W_{t-1}})$ denotes nominal wage inflation, $MRS_t \equiv \int_0^1 MRS_t(h)dh$ gives the average marginal rate of substitution of consumption for leisure, and $RW_t \equiv \int_0^1 \frac{W_t(h)}{P_t}dh$ denotes the average real wage.

The price-inflation equation associated with FS takes the familiar form

$$\pi_t = \beta E_t \pi_{t+1} + \lambda mc_t, \tag{18}$$

where $MC_t \equiv \int_0^1 \frac{MC_t^n(i)}{P_t}di$ denotes the average real marginal cost. In RM, parameter λ takes its standard value $\frac{(1-\beta\theta)(1-\theta)}{\theta}$. Importantly, the determination of the value for parameter λ changes if FS is assumed instead. In that case, its value is computed numerically, as discussed in Woodford (2005). His method posits linearized rules for price setting and for investment:

$$\widehat{p}_t^*(i) = \widehat{p}_t^* - \tau_1 \widehat{k}_t(i), \tag{19}$$

$$\widehat{k}_{t+1}(i) = \tau_2 \widehat{k}_t(i) + \tau_3 \widehat{p}_t(i), \tag{20}$$

where τ_1, τ_2, and τ_3 are parameters that are determined by the method of undetermined coefficients. In stating the decision rules, we have also used the definitions $\widehat{p}_t^*(i) \equiv \log(\frac{P_t^*(i)}{P_t})$, $\widehat{p}_t(i) \equiv \log(\frac{P_t(i)}{P_t})$, $\widehat{p}_t^* \equiv \log(\frac{P_t^*}{P_t})$, and $\widehat{k}_t(i) \equiv \log(\frac{K_t(i)}{K_t})$. For our purposes in the present paper, these rules turn out to be of crucial importance, for they allow us to compute the welfare-relevant second moments of the cross-sectional distributions of prices and capital holdings.

The goods market-clearing equation reflects our assumption that there are subsidies offsetting the distortions associated with monopolistic competition in goods and labor markets. This implies that the steady state of our model is Pareto efficient. Specifically, we have

$$y_t = \zeta c_t + \frac{1-\zeta}{\delta}[k_{t+1} - (1-\delta)k_t], \tag{21}$$

where $\zeta \equiv \frac{\rho+\delta(1-\alpha)}{\rho+\delta}$ denotes the steady-state consumption to output ratio.

The frequency of our model is quarterly. Unless stated otherwise, we assign the following values to the model parameters. We assume $\alpha = 0.36$ for the capital share. Our choice for the value of the risk-aversion parameter σ is 2. The elasticity of substitution between goods, ε, is set to 11, and we assume $\varepsilon_N = 6$ for the elasticity of substitution between different types of labor. Our baseline value for the Calvo parameter for price setting, θ_p, is 0.75, and we assume the same value for its wage-setting counterpart θ_w. The rate of capital depreciation, δ, is assumed to be equal to 0.025, and we set $\epsilon_\psi = 3$ for the capital adjustment cost. These parameter values are justified in Sveen and Weinke (2007) and the references therein. Finally, the coefficient of autocorrelation in the process of technology, ρ_a, is assumed to take the value 0.95, as in Erceg, Henderson, and Levin (2000).

3. Welfare

Let the policymaker's period welfare function be the unweighted average of households' period utility,

$$\mathcal{W}_t \equiv U(C_t) + \int_0^1 V(N_t(h))dh = U(C_t) + E_h\{V(N_t(h))\}. \quad (22)$$

In what follows, period welfare is expressed as a fraction of steady-state Pareto optimal consumption; i.e., we consider $\frac{\mathcal{W}_t - \overline{\mathcal{W}}}{\overline{U}_C \overline{C}}$, where a bar indicates the steady-state value of the original variable and U_C is the marginal utility of consumption. Based on the method in Rotemberg and Woodford (1998), we compute a second-order approximation to period welfare.[12] Our welfare criterion is the unconditional expectation of period welfare, which we write in the way proposed by Svensson (2000):[13]

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

In appendix 1 we derive the following expression for that welfare criterion in the context of FS:

$$\begin{aligned} \widetilde{\mathcal{W}}^{FS} \simeq\ & \Omega_1 E\{y_t^2\} + \Omega_2 E\{c_t^2\} + \Omega_3 E\{i_t^2\} + \Omega_4 E\{(\Delta k_{t+1})^2\} \\ & + \Omega_5 E\{n_t^2\} + \Omega_6 E\{\lambda_t\} + \Omega_1^{FS} E\{\Delta_t\} \\ & + \Omega_2^{FS} E\{\kappa_t\} + \Omega_3^{FS} E\{\psi_t\}, \end{aligned} \quad (23)$$

where the symbol $\simeq$ is meant to indicate that an approximation is accurate up to the second order. The operator E denotes the unconditional expectation. We have also used the following definitions: $\Delta_t \equiv Var_i \widehat{p}_t(i)$, $\kappa_t \equiv Var_i k_t(i)$, $\psi_t \equiv Cov_i(\widehat{p}_t(i), k_t(i))$, and $\lambda_t \equiv Var_h \widehat{w}_t(h)$, with $\widehat{W_t}(h) \equiv \frac{W_t(h)}{W_t}$ and Var_i and Cov_i denoting the cross-sectional variance and covariance operators. Parameters

[12] The proof that the method of Rotemberg and Woodford (1998) can be applied to the problem at hand carries over from Edge's (2003) work to our work because the relevant steady-state properties of the two models are identical.

[13] We scale the discounted sum by $(1-\beta)$ to keep the limit finite as the discount factor goes to unity (from below).

Ω_1 to Ω_6 and Ω_1^{FS} to Ω_3^{FS} are functions of the structural parameters. The latter group of parameters is specific to FS, whereas the former one is common between FS and RM. Both sets of parameters are defined in appendix 1. The main complication that we have to face is to calculate the cross-sectional variances of prices and capital holdings at each point in time as well as their covariance. In order to overcome that difficulty, we make one key observation: Woodford's (2005) linearized rules for price setting and for investment can be used to compute the relevant second moments with the accuracy that is needed for our second-order approximation to welfare.

It is useful to note two properties of equation (23). First, for economically relevant calibrations of our model, the resulting coefficient values in the welfare criterion do not all have the same sign. Second, the variables entering (23) must respect the linearized equilibrium conditions described in the previous section. This is important because these two aspects of (23) are precisely the reason why our welfare criterion does not imply that increasing the variability of those variables that enter with a positive coefficient would lead to an increase in welfare. For instance, increasing the variability of output implies that the associated use of factor inputs would also need to become more variable, and this in turn would reduce welfare.

For the rental-market case, our welfare criterion reads

$$\begin{aligned}\widetilde{\mathcal{W}}^{RM} \simeq\ & \Omega_1 E\{y_t^2\} + \Omega_2 E\{c_t^2\} + \Omega_3 E\{i_t^2\} + \Omega_4 E\{(\Delta k_{t+1})^2\} \\ & + \Omega_5 E\{n_t^2\} + \Omega_6 E\{\lambda_t\} + \Omega^{RM} E\{\Delta_t\}, \qquad (24)\end{aligned}$$

as shown in appendix 2, where we also define parameter Ω^{RM}. Compared with FS, the analysis is greatly simplified by the fact that the capital labor ratio is constant across firms, as discussed in Edge (2003). In the next section we will use our model to analyze the desirability of alternative arrangements for the conduct of monetary policy. This involves a numerical evaluation of the welfare criterion. Before we turn to this, let us make a final remark. It is not necessary for our purposes to rewrite the welfare criterion in the ways proposed by Edge (2003). In order to state the welfare criterion in terms of suitably defined gap terms, she uses the linearized equilibrium conditions to substitute out some of the variables in the welfare measure. This implies cross-terms among the remaining variables, which are absent in our formulation.

4. Results

We consider two prominent families of monetary policy rules and analyze constrained optimal rules; i.e., we restrict attention to a particular subset of possible parameter values that parameterize the rule.[14] Our main objective is to explain how and why the associated constrained optimal values of the policy parameters change in each case depending on whether or not a rental market for capital is assumed. It is useful to note that rational-expectations equilibrium is locally unique (i.e., determinate) under the constrained optimal policies.

4.1 The Welfare Consequences of Responding to Price and Wage Inflation

We start by considering interest rate rules of the form

$$r_t = \rho + \tau_r(r_{t-1} - \rho) + \tau_s[\tau_\omega \omega_t + (1 - \tau_\omega)\pi_t], \tag{25}$$

where parameter τ_s measures the overall responsiveness of the nominal interest rate to changes in inflation, whereas τ_ω is the relative weight put on wage inflation. The weight on price inflation is therefore given by $(1 - \tau_\omega)$. Finally, parameter τ_r denotes the interest rate smoothing coefficient. Only positive parameter values are considered and, moreover, we require parameter τ_ω to be less than or equal to one.[15]

We compare the optimized interest rate rules under FS and RM. In each case we report the optimized coefficients entering the interest

[14] A different approach to welfare-based evaluation of monetary policy is to consider optimal monetary policy. There are, however, different possible and plausible definitions of optimality that have been proposed in the literature. One key distinction is the one between discretion versus commitment, and two important variants of the latter concept are Ramsey optimal policy versus timeless perspective.

[15] Our computational strategy is as follows. First, we define a function that takes the policy parameters as an input and returns the associated value of our welfare criterion. We use DYNARE (http://www.cepremap.cnrs.fr/dynare/) to construct that function. In a second step, we find the optimal parameter configuration among the values we consider. Thanks to Larry Christiano for providing us with Matlab code, which we have used in the computation of λ.

Table 1. Price- and Wage-Inflation Rule

Parameter	FS	RM
τ_r	0.972	0.824
τ_s	2.390	5.852
τ_w	0.415	0.709
Welfare	−9.363	−8.169

rate rule as well as the associated welfare. We follow Erceg, Henderson, and Levin (2000) and isolate the cost of nominal rigidities; i.e., we compute welfare for an economy with flexible prices and flexible wages and subtract the resulting expression from our welfare measure. As they do, we divide this difference by the productivity innovation variance.[16] The results are shown in table 1.

Regardless of whether FS or RM is used, the implied optimized rule prescribes to adjust the nominal interest rate in response to changes in both wage inflation and price inflation. That is intuitive: both kinds of inflation are costly in welfare terms since we model two nominal rigidities. Interestingly, the optimized rule prescribes to react relatively more to price inflation in FS, whereas the opposite holds true in RM. Our intuition is as follows. We observe two things. First, Sveen and Weinke (2005) show that price stickiness can be used to measure the difference between RM and FS, if attention is restricted to a first-order approximation to the equilibrium dynamics: the feature of firm-specific capital implies that price setters internalize the consequences of their price-setting decisions for the marginal cost they face. That makes them more reluctant to change their prices in FS than under RM.[17] Specifically, we show in

[16] Let us give a concrete example for the interpretation of the welfare numbers in our tables. Suppose the productivity innovation variance is 0.01^2. Then, the number -10 for welfare would mean that the representative household would be willing to give up $10 \times 0.01^2 \times 100 = 0.1$ percentage point of steady-state (Pareto optimal) consumption in order to avoid the business-cycle cost associated with the presence of the nominal rigidities in our model.

[17] Similar intuitions have originally been developed by Galí, Gertler, and López-Salido (2001) and Sbordone (2002) in the context of models where capital is assumed to be a constant factor. For an early model featuring differences in the marginal cost across firms, see Woodford (1998).

our 2005 paper that a value of about 0.9 is needed in RM in order to obtain equivalence with FS if the value 0.75 is assigned to the price-stickiness parameter in the latter case and all the remaining parameters are held constant at conventional values. Put differently, the rental-market assumption turns off the *endogenous* price stickiness that is implied by the alternative specification with firm-specific capital. It is an important corollary of our result that an upward-biased estimate of the price-stickiness parameter is obtained if the econometrician looks at the macro data through the lens of RM (if the data-generating process is better described by FS).[18] Second, it is a well-understood property of many New Keynesian models that the central bank achieves the most desirable welfare outcome if it cares relatively more about the nominal variable that is relatively stickier.[19] Combining these two observations, the previous finding seems intuitive. Since the rental-market assumption eliminates the endogenous part of the price stickiness, the central bank should care relatively more about wage inflation in that model. The reason is endogenous wage stickiness. That feature is common to FS and RM: in both models households internalize the consequences of their wage-setting decisions for the marginal disutility of labor they face. On the other hand, if firm-specific capital is taken into account, then the implied endogenous price stickiness is strong enough to make it worthwhile for the central bank to care relatively more about price inflation.

So far our intuition relies on a finding—namely, our price-stickiness metric—that has been obtained in the context of a first-order approximation to the equilibrium dynamics. This intuition could easily be misleading for our purposes here. The reason is that the second-order approximation to the average of households' expected utility, our welfare criterion, is not equivalent in both models if we just change the price stickiness in such a way that the two models would be identical up to the first order. We therefore challenge our intuition by conducting the following experiment, the results of which are shown in table 2.

[18] Based on that observation, Eichenbaum and Fisher (2007) provide some empirical evidence for the plausibility of FS.

[19] See, e.g., Aoki (2001) and Benigno (2003).

Table 2. Robustness I: Rules from the RM Model Used in FS

Parameter	RM Rule with $\theta = 0.75$	RM Rule with $\theta = 0.8947$
τ_r	0.824	1.003
τ_s	5.852	1.095
τ_w	0.709	0.360
Welfare	−12.678	−9.472

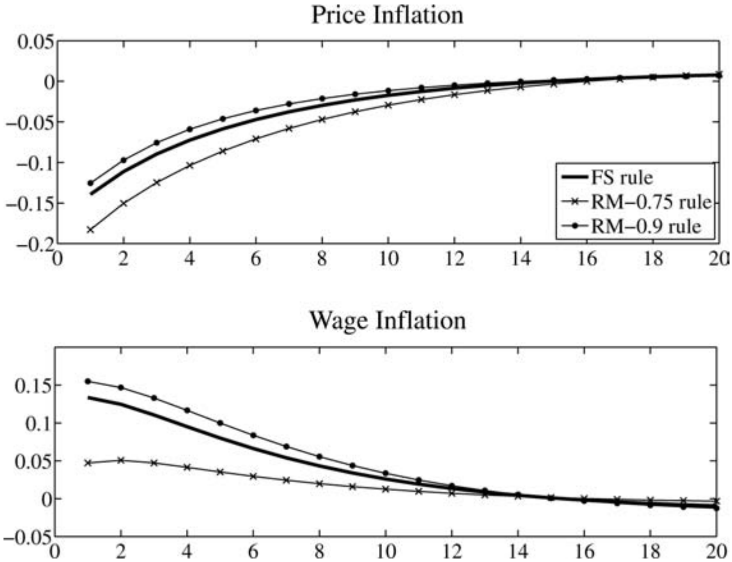
We compute welfare in FS as implied by the optimized policy rule in RM under the baseline calibration. The resulting decrease in welfare is 29.1 percent with respect to the outcome under the optimized rule for FS. Now we compute constrained optimal policy in RM for a price-stickiness parameter equal to 0.8947, in which case RM and FS are identical, up to the first order. The implied optimized rule looks similar to the one associated with FS under the baseline calibration. Specifically, the rule prescribes to react relatively more to price inflation than to wage inflation. Moreover, the decrease in welfare which obtains if that rule is used in FS is just 3.6 percent, which we regard as being negligible. The last result suggests that our price-stickiness metric is useful from a welfare point of view.[20]

To further illustrate the macroeconomic consequences of the three different monetary policy rules, we construct impulse responses to a one-standard-deviation shock to productivity for price inflation and wage inflation. They are shown in figure 1. Under the baseline calibration, the optimal simple rule for FS implies that price inflation is stabilized relatively more than is the case if the optimized rule for RM is used instead. However, if the price-stickiness parameter is set to 0.8947 in RM, then the implied optimized rule delivers an outcome in FS that is similar to the one under the optimized rule for that model.

Next we consider the welfare implications of interest rate rules prescribing that the central bank adjusts the nominal interest rate

[20] In principle, whether or not the price-stickiness metric is useful to tell the difference in welfare implications between FS and RM could depend on the specification of monetary policy. For all the policies we consider, however, our metric turns out to be useful.

Figure 1. Impulse Responses to a Technology Shock with Different Wage-Inflation Rules



not only in response to nominal variables but also as a function of a measure of real economic activity.

4.2 The Welfare Consequences of Taylor-Type Rules

We now turn to the welfare implications of Taylor-type rules,

$$r_t = \rho + \tau_r(r_{t-1} - \rho) + \tau_\pi \pi_t + \tau_y y_t^{gap}, \tag{26}$$

where parameters τ_π and τ_y measure the responsiveness of the nominal interest rate to the respective changes in the rate of price inflation and the output gap. All parameters are assumed to be non-negative. The output gap, y_t^{gap}, is generally defined as the difference between the equilibrium output in an economy with frictions and natural output, i.e., the equilibrium output that would obtain in the absence of nominal frictions. In the context of a model featuring endogenous capital accumulation, Woodford (2003, ch. 5) proposes to refine the notion of natural output in the following way. He uses the equilibrium output that would obtain if the nominal rigidities were absent

Table 3. Taylor-Type Rule with Woodford Output Gap

Parameter	FS	RM
τ_r	1.019	1.641
τ_π	0	0.2618
τ_y	0.096	0.331
Welfare	−9.471	−8.288

and expected to be absent in the future but taking as given the capital stock resulting from optimizing investment behavior in the past in an environment *with* the nominal rigidities present. Woodford argues that this measure of natural output is more closely related to equilibrium determination than the alternative measure which has been used by Neiss and Nelson (2003). Under their definition, natural output is the equilibrium output that would obtain if nominal rigidities were not only currently absent and expected to be absent in the future but had also been absent *in the past*. Indeed, intuitively, the Neiss and Nelson definition of natural output appears to be a bit artificial. We find, however, that from a practical point of view it does not matter for the design of constrained optimal interest rate rules which concept of natural output is used to compute the output gap. We will come back to this point. Before that, let us consider some welfare implications of Taylor-type rules using Woodford's definition of the output gap.[21] The results are shown in table 3.

The optimal rule implied by FS prescribes not to respond to changes in price inflation. On the other hand, under RM, we find that the central bank should react to both price inflation and the output gap. In both models, the optimal interest rate rule features superinertia; i.e., the interest rate smoothing coefficient is larger

[21] Our computational strategy to calculate natural output under Woodford's definition is straightforward. First, we calculate the parameters of the linear function mapping aggregate capital and technology into equilibrium aggregate output in an environment *without* any nominal frictions present. Second, we take the equilibrium value of aggregate capital as implied by FS (or by RM when we study that case), combine it with the level of technology, and compute Woodford's natural output, invoking the above mapping.

Table 4. Robustness II: Rules from the RM Model Used in FS

Parameter	RM Rule with $\theta = 0.75$	RM Rule with $\theta = 0.8947$
τ_r	1.641	1.034
τ_π	0.262	0.059
τ_y	0.331	0.030
Welfare	−9.839	−9.619

than one. Our finding that a strong response to changes in the output gap is desirable on welfare grounds is reminiscent of a result by Erceg, Henderson, and Levin (2000). They show that stabilizing the output gap completely achieves an outcome that is close to the one that is optimal in the context of their model. We have already mentioned the fact that our model is more complicated than theirs since we allow for endogenous capital accumulation. However, the intuition behind their result also seems to be at work in our model: by targeting the output gap, the central bank makes sure that fluctuations in wages and prices take place in such a way that the relatively stickier variable moves less.[22]

Our next result is that the loss is negligible if welfare in FS is computed using the optimized rule implied by RM. We therefore argue that Taylor-type rules are very robust. The results are shown in table 4.

As the last table also indicates, welfare associated with using the rule implied by RM in FS can be further increased if the price stickiness is adjusted in RM in such a way that both models would be identical up to the first order. Once again, our price-stickiness metric turns out to be useful. The policy implications of RM are surprisingly accurate if an upward-biased estimate of the price-stickiness parameter (of the kind that the econometrician actually obtains if she looks at the data through the lens of that model) is used in the analysis. Somewhat surprisingly, however, the optimal relative weight attached to the output gap in RM becomes *smaller* (and

[22]Interestingly, the optimal responsiveness to the output gap is zero if we allow for an output-gap response of the nominal interest rate in addition to the responses to price inflation and wage inflation that we had analyzed before.

Table 5. Taylor-Type Rule with Neiss and Nelson Output Gap

Parameter	FS	RM
τ_r	1.033	2.019
τ_π	0	0.133
τ_y	0.090	0.344
Welfare	−9.416	−8.392

hence less in line with the corresponding value implied by FS) if the price stickiness is increased. That feature appears, however, to be specific to Woodford's definition of natural output, as we are going to see next. Finally, we analyze Taylor-type rules using Neiss and Nelson's (2003) definition of the output gap. The results are reported in table 5.

Overall, optimized rules implied by FS and RM are very similar to the ones obtained before under Woodford's definition of the output gap. In particular, we find again that under RM the optimized rule prescribes to react to both inflation and the output gap, whereas the optimized rule associated with FS features no response to inflation. We also confirm our previous finding that Taylor-type rules are very robust. If the optimized rule implied by RM is used under FS, then the resulting welfare loss is negligible and, moreover, the loss can be further reduced if the price-stickiness parameter is adjusted in RM according to our metric. The results are shown in table 6.

There is only one (small) difference with respect to the previous analysis of Taylor-type rules featuring an output gap à la Woodford. Under the Neiss and Nelson definition, the resulting interest

Table 6. Robustness III: Rules from the RM Model Used in FS

Parameter	RM Rule with $\theta = 0.75$	RM Rule with $\theta = 0.8947$
τ_r	2.019	1.023
τ_π	0.133	0
τ_y	0.344	0.053
Welfare	−9.981	−9.442

rate rules become more similar between FS and RM if we adjust the price stickiness in RM as prescribed by our metric.[23]

Our intuition for why the particular definition of the output gap that is used in the analysis of optimal monetary policy matters so little is simple. The capital stock does not change much at business-cycle frequencies, and the difference between the change in capital implied by a model with and without nominal rigidities present is even less important. In fact, the unconditional correlation between both output-gap measures is 0.9988 in FS under our baseline calibration. This remarkably high correlation between the two variables might, of course, be altered in the presence of additional sources of uncertainty. This caveat notwithstanding, we conclude that the practical relevance of the difference in the two output-gap measures seems to be relatively limited.

Regardless of the definition of the output gap, Taylor-type rules appear to be very robust. The output gap is, of course, not directly observable. However, our results stress the importance of constructing (theory-consistent) observable measures of that variable.

So far, our analysis has been restricted to the welfare comparison of interest rate rules that would be (constrained) optimal under alternative assumptions for capital accumulation. A different question regards the extent to which the welfare properties of a given rule change if we vary its policy parameters. We turn to this next.

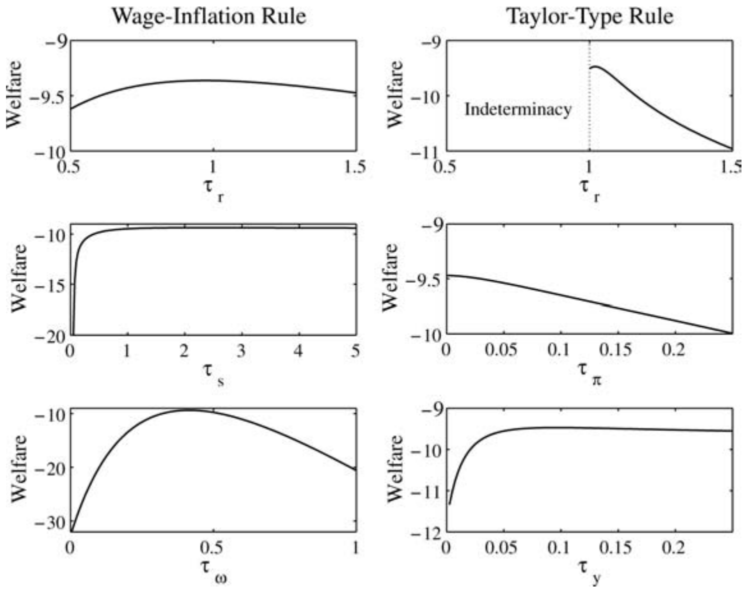
4.3 Sensitivity Analysis

Let us reconsider the Taylor-type rule stated in equation (26). In what follows we restrict attention to the Woodford definition of the output gap.[24] The left-hand panel of figure 2 shows the results of an analogous sensitivity analysis for the case of the wage- and price-inflation interest rate rule in equation (25). Interestingly, we observe that welfare is very sensitive to changes in the parameter that measures the relative importance attached to price and wage inflation in the rule. Also, in this sense, these rules are not robust.

[23]The finding that, if anything, small details of the optimized interest rate rules change depending on which measure of the output gap is used is also confirmed by further robustness checks that we have conducted for alternative interest rate rules.

[24]Additional results are available upon request.

Figure 2. Sensitivity Analysis



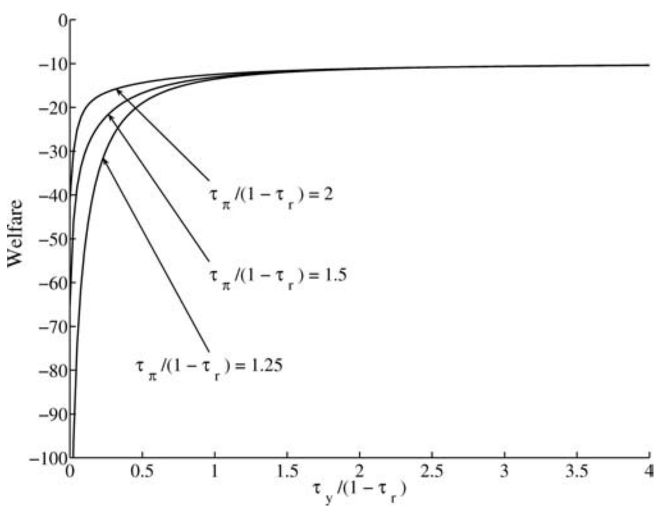
The right-hand panel of figure 2 shows by how much welfare decreases if the policymaker deviates from the optimal rule. Specifically, the figure illustrates the welfare associated with a change in one policy parameter at a time while holding the remaining parameters constant at their (constrained) optimal values. The figure also shows the respective indeterminacy regions.

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

We have already mentioned the fact that constrained optimal Taylor-type rules feature superinertia in both FS and RM. That property does not seem to be realistic. In fact, in estimated central bank reaction functions, the interest rate smoothing coefficient is typically significantly positive but smaller than one (see, e.g., Woodford 2003, ch. 1). It is therefore of interest to examine whether our previous conclusion regarding the robustness of Taylor-type rules remains valid if we restrict attention to an empirically plausible value of the interest rate smoothing coefficient. We set $\tau_r = 0.7$, which is in line with empirical estimates (see, e.g., Galí and Rabanal 2005), and analyze the implications for welfare associated with alternative values for the remaining parameters in equation (26) (including the constrained optimal ones for that case). The results are shown in figure 3.

The last figure makes clear that Taylor-type rules are indeed robust in the sense that we have emphasized before. To the extent that the central bank reacts strongly enough to changes in the output gap, welfare is close to the (constrained) optimal one. Sveen and Weinke (2005, 2007) have made the case for adjusting the nominal interest rate in response to changes in real economic activity based on the argument that the resulting interest rate rules do generally imply determinacy. In the present paper we sharpen our earlier conclusion on welfare grounds. Also, from a welfare point of view, we find that a central bank should react by sufficiently much in response to changes in real economic activity, as measured by the output gap. The last result extends some recent work by Schmitt-Grohé and Uribe (2007b). They show in the context of a New Keynesian

Figure 3. Sensitivity Analysis for Inertial Taylor-Type Rule



model featuring endogenous capital accumulation that large welfare losses are generally implied by policy rules prescribing that the central bank adjusts the nominal interest rate in response to changes in output. We find that a different conclusion obtains if the central bank uses the output gap instead. The welfare properties of an interest rate rule therefore depend crucially on the measure of real economic activity used in its definition.

5. Conclusion

The present paper makes progress in explaining the welfare consequences of firm-specific capital accumulation. We analyze (constrained) optimal interest rate rules prescribing that the nominal interest rate is set as a function of a small number of macroeconomic variables. Our results suggest that Taylor-type interest rate rules are desirable from a welfare point of view.

Appendix 1. Welfare with Firm-Specific Capital

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

$$\frac{A_t - \overline{A}}{\overline{A}} \simeq a_t + \frac{1}{2}a_t^2, \tag{27}$$

where $a_t \equiv \ln(\frac{A_t}{\overline{A}})$. Moreover, if $A_t = \left(\int_0^1 A_t(i)^\gamma di\right)^{\frac{1}{\gamma}}$ then

$$a_t \simeq E_i a_t(i) + \frac{1}{2}\gamma\, Var_i a_t(i). \tag{28}$$

As we have already mentioned in the text, the policymaker's period welfare function reads

$$\mathcal{W}_t \equiv U(C_t) - \int_0^1 V(N_t(h))dh = U(C_t) - E_h\{V(N_t(h))\}. \tag{29}$$

Now we compute a second-order Taylor expansion of period welfare:

$$\mathcal{W}_t \simeq \overline{\mathcal{W}} + \overline{U}_C\overline{C}\left(c_t + \frac{1}{2}c_t^2\right) - \overline{V}_N\overline{N}E_h\left\{n_t(h) + \frac{1}{2}n_t(h)^2\right\} + \frac{1}{2}\overline{U}_{CC}\overline{C}^2c_t^2 - \frac{1}{2}\overline{V}_{NN}\overline{N}^2E_h\{n_t(h)^2\}. \quad (30)$$

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

The Consumption Portion of Welfare

The resource constraint reads

$$Y_t = C_t + I_t. \quad (31)$$

The following relationship holds true:

$$c_t + \frac{1}{2}c_t^2 \simeq \frac{1}{\zeta}\left(y_t + \frac{1}{2}y_t^2\right) - \frac{1-\zeta}{\zeta}\left(i_t + \frac{1}{2}i_t^2\right). \quad (32)$$

Next we analyze the investment portion of the resource constraint. Our starting point is the log-deviation of investment at the firm level:

$$i_t \simeq E_i\{i_t(i)\} + \frac{1}{2}Var_i\{i_t(i)\}. \quad (33)$$

Given our specification of the capital adjustment cost, we have

$$i_t(i) \simeq \frac{1}{\delta}\left[k_{t+1}(i) - (1-\delta)k_t(i) + \frac{1}{2}\left(\varepsilon_\psi - \frac{1-\delta}{\delta}\right)(k_{t+1}(i) - k_t(i))^2\right],$$

$$Var_i\{i_t(i)\} \simeq \frac{1}{\delta^2}[\kappa_{t+1} + (1-\delta)^2\kappa_t - 2(1-\delta)Cov_i\{k_{t+1}(i), k_t(i)\}].$$

Using the investment rule, we derive a second-order approximation to the covariance term in the last relationship:

$$Cov_i\{k_{t+1}(i), k_t(i)\} \simeq \tau_2 Var_i\{k_t(i)\} + \tau_3 Cov_i\{\widetilde{p}_t(i), k_t(i)\}.$$

We also note that

$$E_i\{k_t(i)\} \simeq k_t - \frac{1}{2}\kappa_t.$$

Combining the last five results, we obtain

$$\begin{aligned} i_t \simeq \frac{1}{\delta}\Bigg[&\left(k_{t+1} - \frac{1}{2}\kappa_{t+1}\right) - (1-\delta)\left(k_t - \frac{1}{2}\kappa_t\right) \\ &+ \frac{1}{2}\left(\varepsilon_\psi - \frac{1-\delta}{\delta}\right) E_i\{[k_{t+1}(i) - k_t(i)]^2\}\Bigg] \\ &+ \frac{1}{2}\frac{1}{\delta^2}\kappa_{t+1} + \frac{1}{2}\left(\frac{1-\delta}{\delta}\right)^2 \kappa_t - \frac{1-\delta}{\delta^2}[\tau_2\omega_t + \tau_3\psi_t]. \end{aligned}$$

Let us now consider the term $E_i\{[k_{t+1}(i) - k_t(i)]^2\}$ in the last expression.

$$E_i\{[k_{t+1}(i) - k_t(i)]^2\} \simeq \kappa_{t+1} + k_{t+1}^2 + \kappa_t + k_t^2 - 2E_i\{k_{t+1}(i)k_t(i)\},$$

and $E_i\{k_{t+1}(i)k_t(i)\}$ is approximated with the desired accuracy by

$$E_i\{k_{t+1}(i)k_t(i)\} \simeq k_{t+1}k_t + \tau_2\kappa_t + \tau_3\psi_t.$$

We therefore have

$$\begin{aligned} i_t \simeq\ & \frac{1}{\delta}k_{t+1} - \frac{1-\delta}{\delta}k_t + \frac{1}{2}\frac{1}{\delta}\left(\varepsilon_\psi - \frac{1-\delta}{\delta}\right)(k_{t+1} - k_t)^2 \\ &+ \frac{1}{2}\frac{\varepsilon_\psi}{\delta}\kappa_{t+1} + \frac{1}{2}\frac{\varepsilon_\psi}{\delta}(1 - 2\tau_2)\kappa_t - \frac{\tau_3\varepsilon_\psi}{\delta}\psi_t. \end{aligned} \tag{34}$$

Next we analyze the labor portion of welfare.

The Labor Portion of Welfare

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

$$n_t \simeq E_h n_t(h) + \frac{1}{2}\frac{\zeta - 1}{\zeta} Var_h n_t(h). \tag{35}$$

Aggregate labor demand reads

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

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

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

Invoking the second rule, we obtain

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

and we also note that

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

The last result implies

$$\begin{aligned} E_i b_t(i) &= E_i y_t(i) - x_t - \alpha E_i k_t(i), \\ Var_i b_t(i) &= Var_i y_t(i) + \alpha^2 \kappa_t - 2\alpha Cov_i(y_t(i), k_t(i)). \end{aligned}$$

We therefore obtain

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

From the definitions of K_t and Y_t, it follows that

$$\begin{aligned} E_i k_t(i) &\simeq k_t - \frac{1}{2}\kappa_t, \\ E_i y_t(i) &\simeq y_t - \frac{1}{2}\frac{\varepsilon - 1}{\varepsilon} Var_i y_t(i). \end{aligned}$$

Combining the last three results, we arrive at

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

We therefore obtain

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

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

$$\begin{aligned}Var_i y_t(i) &\simeq \varepsilon^2 \Delta_t, \\ Cov_i(y_t(i), k_t(i)) &\simeq -\varepsilon \psi_t, \\ Var_h n_t(h) &\simeq \varepsilon_N^2 \lambda_t.\end{aligned}$$

We therefore have

$$\begin{aligned}E_h n_t(h) \simeq{}& \frac{1}{1-\alpha}(y_t - x_t - \alpha k_t) + \frac{1}{2}\frac{1-\alpha+\alpha\varepsilon}{(1-\alpha)^2}\varepsilon\Delta_t \\ &+ \frac{1}{2}\frac{\alpha}{(1-\alpha)^2}\kappa_t + \frac{\alpha\varepsilon}{(1-\alpha)^2}\psi_t - \frac{1}{2}(\varepsilon_N - 1)\varepsilon_N\lambda_t. \qquad (36)\end{aligned}$$

Finally, we note that $E_h\{n_t(h)^2\}$ can be written as

$$E_h\{n_t(h)^2\} \simeq \varepsilon_N^2 \lambda_t + n_t^2. \qquad (37)$$

The Welfare Function

The Pareto optimality of the steady state implies $\overline{V}_N \overline{N} = U_C \overline{C}\frac{1-\alpha}{\zeta}$. Combining that result with equations (30), (32), (34), (36), and (37), we arrive at the following approximation to period welfare:

$$\begin{aligned}\frac{\mathcal{W}_t-\overline{\mathcal{W}}}{\overline{U}_C\overline{C}} \simeq{}& \frac{1}{\zeta}x_t-\frac{1}{\zeta}\frac{\alpha}{\rho+\delta}[k_{t+1}-(1+\rho)k_t]+\frac{1}{2}\frac{1}{\zeta}y_t^2-\frac{\sigma}{2}c_t^2\\ &-\frac{1}{2}\frac{1-\zeta}{\zeta}i_t^2-\frac{1}{2}\frac{1-\zeta}{\zeta}\frac{1}{\delta}\left(\varepsilon_\psi-\frac{1-\delta}{\delta}\right)(k_{t+1}-k_t)^2\\ &-\frac{1}{2}\frac{1-\alpha}{\zeta}(1+\phi)n_t^2-\frac{1}{2}\frac{\varepsilon}{\zeta}\frac{1-\alpha+\alpha\varepsilon}{1-\alpha}\Delta_t\\ &-\frac{1}{2}\frac{(1-\alpha)\varepsilon_N}{\zeta}(1+\varepsilon_N\phi)\lambda_t-\frac{1}{2}\frac{1-\zeta}{\zeta}\frac{\varepsilon_\psi}{\delta}\kappa_{t+1}\\ &-\frac{1}{2}\frac{1}{\zeta}\left[\frac{(1-\zeta)\varepsilon_\psi}{\delta}(1-2\tau_2)+\frac{\alpha}{1-\alpha}\right]\kappa_t\\ &-\frac{1}{\zeta}\left[\frac{\alpha\varepsilon}{1-\alpha}-\frac{(1-\zeta)\tau_3\varepsilon_\psi}{\delta}\right]\psi_t. \end{aligned}\tag{38}$$

The first linear term in the last expression is proportional to the level of aggregate (log) technology, x_t, which is exogenous. The remaining linear terms are proportional to current and next period's aggregate capital, k_t and k_{t+1}. Next we compute $\frac{1-\beta}{\overline{U}_C\overline{C}}E_t\{\sum_{k=0}^{\infty}\beta^k(\mathcal{W}_{t+k}-\overline{\mathcal{W}})\}$, which allows us to invoke a result by Edge (2003). As in her model, the terms in aggregate capital cancel except for the initial one. Following the lead of Svensson (2000), we consider the limit for $\beta\to 1$. This allows us to abstract from initial conditions.[25] This gives the expression stated in the text:

$$\begin{aligned}\widetilde{\mathcal{W}}^{FS} \simeq{}& \Omega_1 E\{y_t^2\}+\Omega_2 E\{c_t^2\}+\Omega_3 E\{i_t^2\}+\Omega_4 E\{(\Delta k_{t+1})^2\}\\ &+\Omega_5 E\{n_t^2\}+\Omega_6 E\{\lambda_t\}+\Omega_1^{FS}E\{\Delta_t\}\\ &+\Omega_2^{FS}E\{\kappa_t\}+\Omega_3^{FS}E\{\psi_t\}, \end{aligned}\tag{39}$$

with

$$\begin{aligned}&\Omega_1\equiv\frac{1}{2}\frac{1}{\zeta},\ \Omega_2\equiv-\frac{\sigma}{2},\Omega_3\equiv-\frac{1}{2}\frac{1-\zeta}{\zeta},\\ &\Omega_4\equiv-\frac{1}{2}\frac{1}{\delta}\frac{1-\zeta}{\zeta}\left(\varepsilon_\psi-\frac{1-\delta}{\delta}\right),\end{aligned}$$

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

$$\Omega_5 \equiv -\frac{1}{2}\frac{1-\alpha}{\zeta}(1+\phi), \Omega_6 \equiv -\frac{1}{2}\frac{(1-\alpha)\varepsilon_N}{\zeta}(1+\phi\varepsilon_N),$$
$$\Omega_1^{FS} \equiv -\frac{1}{2}\frac{\varepsilon}{\zeta}\frac{1-\alpha+\alpha\varepsilon}{1-\alpha},$$
$$\Omega_2^{FS} \equiv -\frac{1}{2}\frac{1}{\zeta}\left\{\frac{2(1-\zeta)\varepsilon_\psi(1-\tau_2)}{\delta}+\frac{\alpha}{1-\alpha}-\frac{2}{\delta}\frac{1-\zeta}{\delta}\right\},$$
$$\Omega_3^{FS} \equiv -\frac{1}{\zeta}\left(\frac{\alpha\varepsilon}{1-\alpha}-\frac{(1-\zeta)\tau_3\varepsilon_\psi}{\delta}\right).$$

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

$$\lambda_t \simeq \theta_w\lambda_{t-1}+\frac{\theta_w}{1-\theta_w}(\omega_t)^2. \tag{40}$$

We define $\widetilde{p}_t \equiv E_i \ln P_t(i)$ and write the price-dispersion term, Δ_t, in the way proposed by Woodford (2003, ch. 6):

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

After invoking the price-setting rule stated in equation (19), we obtain the following recursive formulation for Δ_t:

$$\begin{aligned}\Delta_t &\simeq \theta_p\Delta_{t-1}+(1-\theta_p)E_i\left[\left(\widetilde{p}_t^*-\tau_1\widehat{k}_t(i)-\widetilde{p}_{t-1}\right)^2\right]-(\widetilde{p}_t-\widetilde{p}_{t-1})^2\\ &\simeq \theta_p\Delta_{t-1}+(1-\theta_p)E_i\left[\left(\frac{1}{1-\theta_p}(\widetilde{p}_t-\widetilde{p}_{t-1})-\tau_1\widehat{k}_t(i)\right)^2\right]\\ &\quad -(\widetilde{p}_t-\widetilde{p}_{t-1})^2\\ &\simeq \theta_p\Delta_{t-1}+(1-\theta_p)\tau_1^2\kappa_t+\frac{\theta_p}{1-\theta_p}\pi_t^2.\end{aligned} \tag{41}$$

The investment rule stated in equation (20) implies directly that a recursive formulation for the cross-sectional variance of capital holdings, κ_t, is of the form

$$\kappa_t \simeq \tau_2^2 \kappa_{t-1} + \tau_3^2 \Delta_{t-1}. \tag{42}$$

Finally, the cross-sectional covariance of prices and capital holdings, ψ_t, can also be expressed in a recursive way:

$$\begin{aligned}
\psi_t \simeq\ & E_i[\widehat{k}_t(i)\widehat{p}_t(i)], \\
\simeq\ & \theta_p E_i[\widehat{k}_t(i)(\ln P_{t-1}(i) - \ln P_t)] + (1-\theta_p)E_i\left[\widehat{k}_t(i)\widehat{p}_t^*(i)\right], \\
\simeq\ & \theta_p E_i[\left(\tau_2 \widehat{k}_{t-1}(i) + \tau_3 \widehat{p}_{t-1}(i)\right)\widehat{p}_{t-1}(i)] \\
& + (1-\theta_p)E_i\left[\widehat{k}_t(i)\left(\widehat{p}_t^* - \tau_1 \widehat{k}_t(i)\right)\right], \\
\simeq\ & \theta_p \tau_2 \psi_{t-1} + \theta_p \tau_3 \Delta_{t-1} - \tau_1(1-\theta_p)\kappa_t.
\end{aligned} \tag{43}$$

Appendix 2. Welfare with Rental Market

In the rental-market case, the analysis is simplified by that fact that the capital labor ratio is constant across firms, as discussed in Edge (2003). The resulting welfare criterion reads

$$\begin{aligned}
\widetilde{\mathcal{W}}^{RM} \simeq\ & \Omega_1 E\left\{y_t^2\right\} + \Omega_2 E\left\{c_t^2\right\} + \Omega_3 E\left\{i_t^2\right\} + \Omega_4 E\{(\Delta k_{t+1})^2\} \\
& + \Omega_5 E\left\{n_t^2\right\} + \Omega_6 E\{\lambda_t\} + \Omega^{RM} E\{\Delta_t\},
\end{aligned} \tag{44}$$

with $\Omega^{RM} \equiv -\frac{1}{2}\frac{\varepsilon}{\zeta}$. Finally, the variance terms can be written as

$$\Delta_t = \theta_p \Delta_{t-1} + \frac{\theta_p}{1-\theta_p}\pi_t^2, \tag{45}$$

$$\lambda_t = \theta_w \lambda_{t-1} + \frac{\theta_w}{1-\theta_w}(\Delta w_t)^2, \tag{46}$$

as discussed in Erceg, Henderson, and Levin (2000) and Rotemberg and Woodford (1999).

References

Altig, D., L. J. Christiano, M. Eichenbaum, and J. Lindé. 2005. "Firm-Specific Capital, Nominal Rigidities, and the Business Cycle." NBER Working Paper No. 11034.

Aoki, K. 2001. "Optimal Monetary Policy Responses to Relative-Price Changes." *Journal of Monetary Economics* 48 (1): 55–80.

Benigno, P. 2003. "Optimal Monetary Policy in a Currency Area." *Journal of International Economics* 63 (2): 293–320.

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

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

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

Eichenbaum, M., and J. D. M. Fisher. 2007. "Estimating the Frequency of Price Re-optimization in Calvo-Style Models." *Journal of Monetary Economics* 54 (7): 2032–47.

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

Galí, J., M. Gertler, and J. D. López-Salido. 2001. "European Inflation Dynamics." *European Economic Review* 45 (7): 1237–70.

Galí, J., and P. Rabanal. 2005. "Technology Shocks and Aggregate Fluctuations: How Well Does the RBC Model Fit Postwar U.S. Data?" In *NBER Macroeconomics Annual 2004*, ed. M. Gertler and K. Rogoff, 225–88. Cambridge, MA: MIT Press.

Levin, A. T., A. Onatski, J. C. Williams, and N. Williams. 2006. "Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models." In *NBER Macroeconomics Annual 2005*, ed. M. Gertler and K. Rogoff, 229–87. Cambridge, MA: MIT Press.

Levin, A. T., V. Wieland, and J. C. Williams. 1999. "Robustness of Simple Monetary Policy Rules under Model Uncertainty." In *Monetary Policy Rules*, ed. J. B. Taylor, 263–99. Chicago: NBER and University of Chicago Press.

Levin, A. T., and J. C. Williams. 2003. "Robust Monetary Policy with Competing Reference Models." *Journal of Monetary Economics* 50 (5): 945–75.

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

Rotemberg, J. J., and M. Woodford. 1998. "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy." In *NBER Macroeconomics Annual 1997*, ed. B. S. Bernanke and J. J. Rotemberg, 297–346. Cambridge, MA: MIT Press.

———. 1999. "Interest Rate Rules in an Estimated Sticky Price Model." In *Monetary Policy Rules*, ed. J. B. Taylor, 57–119. Chicago: NBER and University of Chicago Press.

Sbordone, A. M. 2002. "Prices and Unit Labor Costs: A New Test of Price Stickiness." *Journal of Monetary Economics* 49 (2): 265–92.

Schmitt-Grohé, S., and M. Uribe. 2006. "Optimal Fiscal and Monetary Policy in a Medium-Scale Macroeconomic Model." In *NBER Macroeconomics Annual 2005*, ed. M. Gertler and K. Rogoff, 83–425. Cambridge, MA: MIT Press.

———. 2007a. "Optimal Inflation Stabilization in a Medium-Scale Macroeconomic Model." In *Monetary Policy Under Inflation Targeting*, ed. K. Schmidt-Hebbel and F. S. Mishkin, 125–86. Santiago, Chile: Central Bank of Chile.

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

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

———. 2007. "Firm-Specific Capital, Nominal Rigidities and the Taylor Principle." *Journal of Economic Theory* 136 (1): 729–37.

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

Woodford, M. 1998. "Control of Public Debt: A Requirement for Price Stability?" In *The Debt Burden and Its Consequences for Monetary Policy*, ed. G. A. Calvo and M. King, 117–54. (Proceedings of a conference held by the International Economic Association at the Deutsche Bundesbank, Frankfurt, Germany.) London: Palgrave Macmillan.

———. 2003. *Interest & Prices: Foundations of a Theory of Monetary Policy.* Princeton, NJ: Princeton University Press.

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

Instability and Nonlinearity in the Euro-Area Phillips Curve*

Alberto Musso[a], Livio Stracca[a], and Dick van Dijk[b]
[a]European Central Bank
[b]Econometric Institute, Erasmus University Rotterdam

This paper provides a comprehensive analysis of the functional form of the euro-area Phillips curve over the past three decades. In particular, compared with previous literature, we analyze the stability of the relationship in detail, especially as regards the possibility of a time-varying mean of inflation. Moreover, we conduct a sensitivity analysis across different measures of economic slack. Our main findings are two. First, there is strong evidence of time variation in the mean and slope of the Phillips curve occurring in the early to mid-1980s, but not in inflation persistence once the mean shift is allowed for. As a result of the structural change, the Phillips curve became flatter around a lower mean of inflation. Second, we find no significant evidence of nonlinearity—in particular, in relation to the output gap.

JEL Codes: E52, E58.

1. Introduction

The dynamics of inflation have changed substantially in many, if not all, advanced economies over the past four decades. For example, the average level of inflation has been subject to dramatic shifts over time (Cecchetti et al. 2007). Moreover, in recent years a number

*The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the European Central Bank. We thank Geoff Kenny and Jarkko Turunen, and participants at an ECB seminar, at the workshop on "Nonlinear Dynamical Methods and Time Series Analysis" held in Udine, Italy, August 2006, and at the "2008 conference of the Society for Nonlinear Dynamics and Econometrics" held at the Federal Reserve Bank of San Francisco, April 2008, for useful comments and suggestions. All remaining errors are our own responsibility. Author e-mails: alberto.musso@ecb.int (corresponding author), livio.stracca@ecb.int, djvandijk@few.eur.nl.

of studies have documented important changes also in the degree of inflation persistence (Cecchetti and Debelle 2006). In addition, the volatility of inflation has changed during the past three decades, with a large decline observed since the mid-1980s to early 1990s, depending on the country (van Dijk, Osborn, and Sensier 2002). As a result of these changes, modeling and forecasting inflation dynamics has become an arduous task. The complexity in modeling inflation dynamics relates not only to the various types of above-mentioned structural changes in the statistical properties of inflation, but also to the fact that to some extent these changes are related to one another in various ways. For example, a key result of the Eurosystem Inflation Persistence Network (IPN) is that estimates of the euro-area inflation persistence tend to be rather high unless shifts in the mean of inflation (for which there is clear statistical evidence) are allowed for (Altissimo, Ehrmann, and Smets 2006). Hence, it is important to analyze these changes jointly. At the same time, modeling inflation is complicated also by the fact that in addition to its (potential) instability, different forms of nonlinearity can be relevant. For example, some studies have pointed out the possibility that the response of inflation to changes in economic activity may be asymmetric, with demand increases having a stronger impact on prices than demand decreases (Laxton, Rose, and Tambakis 1999).

While much effort has been devoted to analyzing the inflation process for the U.S. economy, much less research has been undertaken for the euro area. As a result, it is still uncertain how to best model euro-area inflation. This gap is rather unfortunate, given the mandate of the European Central Bank (ECB), whose primary objective is to ensure price stability at the euro-area aggregate level. Some efforts have been directed in recent years to analyzing euro-area inflation dynamics, especially in the context of the so-called New Keynesian Phillips curve (NKPC) framework, with mixed results, as discussed in more detail below. As regards the traditional Phillips-curve approach, relatively little has been done to assess its usefulness for the euro area. The few existing studies, such as Aguiar and Martins (2005), Dolado, Maria-Dolores, and Naveira (2005), Rudd and Whelan (2005), and Baghli, Cahn, and Fraisse (2007), include only a limited analysis of possible instability and nonlinearities. As a result, several questions remained unanswered regarding the most appropriate way to model inflation dynamics in the euro area.

Although an assessment of the functional form of the Phillips curve is fraught with empirical difficulties, the policy implications of this question are extremely important. Let us consider, for example, the situation of a policymaker who is uncertain as to whether the Phillips curve has a linear or, alternatively, a piecewise linear form as in Filardo (1998) and Barnes and Olivei (2003). In the first case, the policymaker is confronted with a trade-off between stimulating demand and creating inflation, while in the latter case there is the possibility of pushing demand at least up to a certain limit without causing a significant increase in inflation. Therefore, a careful empirical modeling of the functional form of the Phillips curve is of paramount importance.

Against this background, the aim of the present paper is to provide a comprehensive analysis of euro-area inflation dynamics, focusing on the functional form of the Phillips curve. We explicitly and carefully address the stability of the relationship between inflation and economic activity, accounting for the possibility of structural change in the mean, persistence, and volatility of inflation, as well as in the slope of the curve. In addition, we examine the appropriate functional form of the curve by means of the methodology of smooth transition regression (STR) models, which allows for both convex and concave shapes of the curve. Although our main analysis is conducted on quarterly inflation based on the GDP deflator, we also analyze the price index that is preferred by the ECB, the Harmonised Index of Consumer Prices (HICP). For the latter indicator, we also analyze the possible presence of nonlinearity in the effect of additive price shocks stemming from oil and exchange rate developments. Finally, we conduct a thorough sensitivity analysis across different possible measures of economic slack.

The paper is structured as follows. In section 2 we present a review of the literature. The data are described in section 3. Section 4 presents the results for a linear Phillips-curve specification for the euro-area GDP deflator inflation. In section 5 we assess the stability and linearity of this curve. In section 6 we model the HICP inflation rate indirectly, by modeling the spread between the HICP and the GDP deflator. Finally, section 7 concludes.

2. Literature Review and Modeling Issues

2.1 Inflation Modeling

The focus of this paper is on the general class of traditional backward-looking Phillips curves. This choice is suggested by a number of considerations. First, survey-based inflation-forecast data for the euro area starting from the 1970s are not available. Second, alternative estimation approaches based on the generalized method of moments which abstract from inflation forecasts are surrounded by a number of controversial econometric aspects, limiting the reliability of NKPC estimates. Third, recent studies—in particular, by Rudd and Whelan (2007)—cast doubt on the ability of the NKPC (including its hybrid form, i.e., with added lags of inflation) to provide a useful empirical characterization of the inflation process and present evidence in support of the traditional Phillips curve for both the United States and the euro area. While we do not take a stand on this debate, we note that it makes the estimation of a backward-looking Phillips curve at least not a clearly suboptimal choice. Finally, it should be emphasized that we conduct a thorough stability analysis in this paper, and in so doing we cater for the possible impact of the Lucas critique, which is often mentioned as the main shortcoming of backward-looking macroeconomic models.

Although traditional Phillips-curve relationships are building blocks of a number of macroeconomic models for the euro area, including the area-wide model (AWM), relatively few studies have provided a detailed modeling assessment of this key relationship. A number of studies providing estimates of the traditional Phillips curve in the euro area have been published in recent years.[1] However, no consensus seems to prevail as regards the most appropriate specification of the relationship. For example, Dolado, Maria-Dolores, and Naveira (2005) and Baghli, Cahn, and Fraisse (2007) provide some evidence for the relevance of nonlinearity in the euro-area Phillips curve, while Aguiar and Martins (2005) suggest that the empirical evidence against the linear specification is weak. Rudd and Whelan (2005) do not consider nonlinear specifications but

[1] Other papers include a euro-area Phillips curve as a component of a broader, multivariate framework, such as Rünstler (2002), Fabiani and Mestre (2004), Fagan, Henry, and Mestre (2005), and Proietti, Musso, and Westermann (2007).

conduct an extensive stability analysis of the linear Phillips relationship and find little evidence of instability. The main reasons for these contrasting results can be related to different sample periods and different specifications, but data issues also may play a role. In particular, the measures used for capturing economic slack tend to differ and range from output-gap estimates based on simple filters (Dolado, Maria-Dolores, and Naveira 2005; Rudd and Whelan 2005) to estimates based on more structural unobserved-components models (Aguiar and Martins 2005; Baghli, Cahn, and Fraisse 2007). Sensitivity analysis to assess how results vary using alternative slack estimates is typically very limited or even missing in these studies. Given the uncertainty surrounding these estimates, this could turn out to be a significant limitation.

2.2 Instability

As discussed in the Introduction, various forms of instability in inflation dynamics have been documented for most advanced economies, including structural changes in the mean, persistence, and volatility of inflation. Focusing on the euro area, Corvoisier and Mojon (2005) find three breaks for the euro-area inflation rate: in 1972 and 1985 with reference to the CPI/HICP, and in 1993 using the GDP deflator. Angeloni, Aucremanne, and Ciccarelli (2006) present evidence of a permanent decline in the persistence of inflation in the euro area after the mid-1990s, even after allowing for breaks in the mean of inflation.

While it is important to take into account these instabilities, it is questionable whether the most appropriate way to detect and model them is via structural-break tests assuming *abrupt* changes. In particular, consistent with the idea that most regime changes tend to be gradual, several studies (especially on the United States) adopt modeling approaches based on smoothly time-varying coefficients, rather than assuming abrupt changes (see, for example, Stock and Watson 2007). We follow this suggestion here by adopting the smooth transition regression framework.

Several papers have found evidence of instability also in the slope of the Phillips curve, i.e., on the response of inflation to demand pressures. In particular, some studies have highlighted the possibility that the Phillips curve may have flattened—i.e., the slope may have decreased—in several advanced economies (Borio and Filardo 2007). The interpretation of this change in the slope of the Phillips curve is still an open issue. A hypothesis which has received much attention is that the source of this flattening may be related to the process of globalization (Melick and Galati 2006). Other authors, such as Roberts (2006), attribute the reduction in the slope to changes in monetary policy. However, there does not seem to be robust evidence for this hypothesis, as recently shown by Ihrig et al. (2007).

Some evidence for significant changes over time also has been uncovered with regard to the impact of oil and exchange rate shocks to inflation. For example, a number of studies have documented a significant decline in the pass-through of oil prices to consumer price inflation in several advanced economies since the 1980s (De Gregorio, Landerretche, and Neilson 2007). Blanchard and Galí (2007) confirm this finding and conclude that various forces have caused this decline, including improved monetary policy, more flexible labor markets, and a smaller dependence on oil. Other studies have provided evidence for a reduced exchange rate pass-through to consumer price inflation in advanced economies after the 1980s, although this decline is not always statistically significant (Ihrig, Marazzi, and Rothenberg 2006).

2.3 Nonlinearity and Asymmetry

There is a long tradition of thought in monetary economics, going back at least to the times of John Maynard Keynes, suggesting that the Phillips curve may be nonlinear and, in particular, have a *convex* shape, reflecting the existence of discontinuity in firms' price adjustment costs—for example, due to downward wage rigidity (e.g., Clark and Laxton 1997). A convex Phillips curve implies that inflation may fail to decline in response to a shortfall of excess demand but pick up significantly should demand exceed a certain threshold: the marginal reaction of inflation to a spending stimulus—for example, coming from monetary policy—is therefore path dependent. An extreme form of convexity is an *asymmetric* curve, where inflation reacts to excess demand only if the latter is above a certain level. It is worth noting that, in fact, the relationship initially proposed by Phillips was, indeed, a curve.

The existing empirical evidence for the United States and other industrialized economies is, however, mixed. Akerlof, Dickens, and Perry (1996) and Debelle and Laxton (1997), among others, suggest that a convex Phillips curve is appropriate, while Gordon (1997) argues in favor of a linear curve and Stiglitz (1997) even of a concave one. The evidence on the functional form of the Phillips curve is particularly scant and controversial in the euro area, partly reflecting the challenges associated with gathering appropriately harmonized and long time series of data for this economy compared, for example, with the United States. Interestingly, research conducted within the Eurosystem IPN has found that prices in the euro area appear to respond more strongly to cost increases than to decreases but, at the same time, more to a fall in demand than to a rise (Fabiani et al. 2006). Transposing this micro evidence to the macroeconomic level, the first bit of evidence would point to a convex Phillips curve, while the second bit suggests a concave curve. On the whole, therefore, the IPN evidence does suggest the existence of some interesting nonlinearity, but the implications at the aggregate level are unclear.

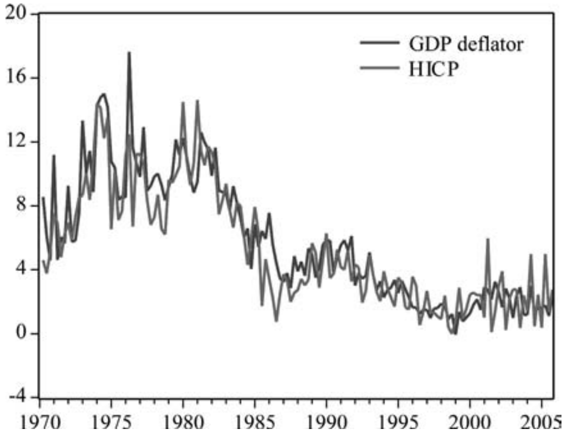
Aguiar and Martins (2005) test the linearity of the euro-area Phillips curve using data from 1970 to 2002 and find that there is not enough statistical evidence for rejecting the null of linearity. However, Dolado, Maria-Dolores, and Naveira (2005) suggest that nonlinearities may be present, working on data from 1984 to 2001. In particular, in their specification, the square value of the output gap enters significantly and with a positive coefficient in the equation, suggesting a convex Phillips curve.

3. Data

The data for our empirical analysis is obtained from the area-wide model (AWM) database[2] and has quarterly frequency, spanning the period 1970:Q1–2005:Q4. We focus on the two main measures of inflation for the euro area, which are based on the GDP deflator and the Harmonised Index of Consumer Prices (HICP). Although the latter is the main indicator referred to by the ECB and is the

[2]For more details on the AWM database, see Fagan, Henry, and Mestre (2005). We make use of the database version released in September 2006, which extends through 2005:Q4.

Figure 1. GDP Deflator and HICP Inflation



Note: The graph shows annualized quarter-on-quarter inflation rates for the euro-area GDP deflator and HICP for the period 1970:Q2–2005:Q4.

ultimate target of our analysis, as discussed below there are certain benefits in starting from a model for the former and then analyzing the link between these two price series.[3] Figure 1 shows the developments of the GDP deflator and the HICP over the sample period in terms of annualized quarter-on-quarter inflation rates. Although the two series move closely together[4] and follow broadly similar patterns, sizable deviations can be observed over some prolonged periods such as the late-1970s and mid-1980s. Moreover, while the GDP deflator also is available in seasonally adjusted form, the HICP only comes in seasonally unadjusted form, a fact that has to be borne in mind in the modeling process.

Typical measures of the output gap are surrounded by a large degree of uncertainty; see Camba-Méndez and Rodriguez-Palenzuela (2003) and Orphanides and van Norden (2005), among others. For that reason, we consider several alternative indicators of economic slack. First, we employ three alternative estimates of the output gap based on the multivariate unobserved-components model of

[3]This approach is frequently adopted in several macroeconomic models that specify a Phillips-type relationship, including the AWM.

[4]Over the complete sample period, the correlation between GDP deflator and HICP inflation is 0.89.

Proietti, Musso, and Westermann (2007).[5] Second, we use three frequently used measures based on statistical filters applied to real GDP: the Baxter-King band-pass filter, the Hodrick-Prescott filter, and a univariate unobserved-components model.[6]

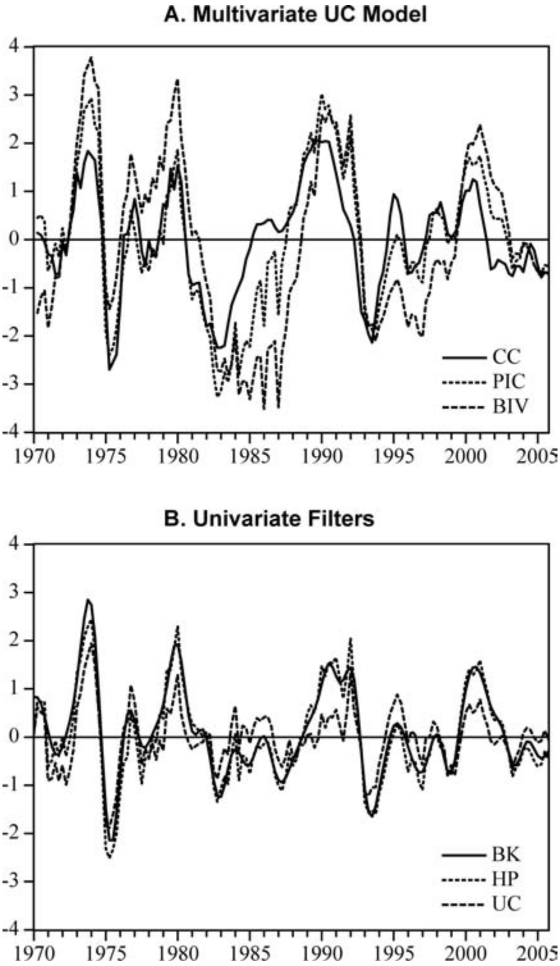
Figure 2 displays the output-gap measures that we consider, while table 1 reports summary statistics. From the graph it appears that although all six variables are highly correlated, their amplitude tends to vary. The large positive cross-correlations in table 1 confirm that there is a great deal of co-movement across the different output-gap measures. At the same time, it is also clear that there is no perfect collinearity among them. To avoid the peculiarities of a specific output-gap measure, in the empirical analysis in the following sections we will make use of their first principal component, which is shown in figure 3. In section 5.1 we conduct a sensitivity analysis where we consider the individual output-gap estimates and an alternative summary measure based on their simple average.

In part of the analysis, two common indicators of additive price shocks also are included—namely, the quarter-on-quarter growth rates in the euro nominal effective exchange rate (standardized to equal 100 in 1970:Q1) and the price of oil (in euros per barrel). The levels of these variables are plotted in figure 4. The nominal effective exchange rate largely resembles the movements of European currencies against the U.S. dollar, reaching a low value of 82

[5]The three versions of the multivariate unobserved-components model, based on the production function approach, consist of the common cycles (CC) version, the pseudo-integrated cycles (PIC) version, and the bivariate (BIV) version. The CC specification is estimated under the assumption that all cyclical variables in the system (total factor productivity, unemployment, and labor force participation) follow the relatively short cycle in capacity utilization. The PIC specification is estimated under the assumption that the cycles in the labor variables are more persistent. The BIV specification is based on a bivariate system for inflation and output only. See Proietti, Musso, and Westermann (2007) for more details.

[6]The three univariate filters are applied to real GDP, extended backwards (using a euro-area aggregate based on OECD data) and projected forwards (with a simple autoregressive model) by three years. Subsequently, the first and last three years of the estimated cycles were discarded, as recommended by Baxter and King (1999) for the band-pass filter. The univariate unobserved-components model was specified as a basic smooth unobserved-components model (fixed level, stochastic slope) with a stochastic cycle (with damping factor equal to 0.9 and period equal to 20) and outlier corrections (found via tests based on the auxiliary residuals) in 1974:Q3, 1986:Q1, and 1987:Q1.

Figure 2. Output-Gap Measures



Note: The graphs show measures of the quarterly output gap. In panel A, CC, PIC, and BIV denote measures obtained from the common cycles, the pseudo-integrated cycles, and the bivariate versions, respectively, of the multivariate unobserved-components model of Proietti, Musso, and Westermann (2007). In panel B, BK denotes the Baxter-King band-pass filter, HP the Hodrick-Prescott filter, and UC a univariate unobserved-components model applied to quarterly read GDP.

in 1985, followed by a rapid increase due to the Plaza Agreement, then a substantial depreciation following the introduction of the euro in 1999, and with the subsequent recovery during 2001–03. The oil

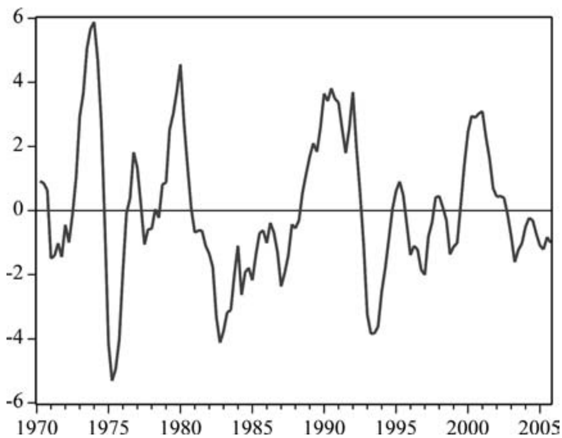
Table 1. Output-Gap Measures—Summary Statistics

					Correlation					
	Mean	**St. Dev.**	**Skewness**	**Kurtosis**	**CC**	**PIC**	**BIV**	**BK**	**HP**	**UC**
CC	−0.01	1.07	−0.22	2.78		0.87	0.61	0.75	0.74	0.63
PIC	−0.03	1.42	0.03	2.62			0.83	0.81	0.82	0.59
BIV	−0.06	1.76	0.15	2.16				0.82	0.79	0.52
BK	0.11	0.92	0.43	3.39					0.97	0.81
HP	0.02	0.97	0.25	3.06						0.83
UC	−0.01	0.63	−0.01	4.07						

Note: The table presents summary statistics for quarterly output-gap measures for the euro area for the period 1970:Q1–2005:Q4. CC, PIC, and BIV are obtained from the common cycles, the pseudo-integrated cycles, and the bivariate versions, respectively, of the multivariate unobserved-components model of Proietti, Musso, and Westermann (2007). BK denotes the Baxter-King band-pass filter, HP the Hodrick-Prescott filter, and UC a univariate unobserved-components model applied to quarterly real GDP.

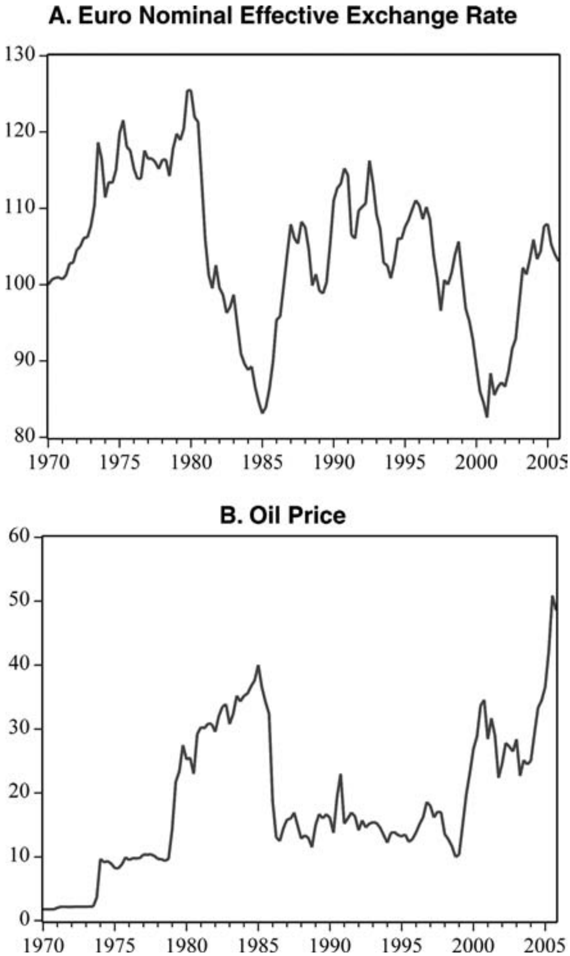
price clearly shows the OPEC-induced price jumps in 1973 and 1979, the rapid decline in 1985–86 following the increase in production initiated by Saudi Arabia, and the price hikes around the turn of the millennium and in 2004–05.

Figure 3. Principal Component of Output-Gap Measures



Note: The graph shows the first principal component of the six output-gap measures for the period 1970:Q1–2005:Q4.

Figure 4. Price Shocks: Euro Nominal Effective Exchange Rate and Oil Price



Note: The graphs show the quarterly euro nominal effective exchange rate and oil price for the period 1970:Q1–2005:Q4.

4. Linear Phillips-Curve Specification

The main conclusion that we draw from the literature review in section 2 is that a comprehensive modeling strategy is required in order to discriminate among alternative specifications for euro-area inflation dynamics and the Phillips curve—in particular, to account for the possible presence of various types of instabilities and nonlinearity. We start from a generalized form of the Phillips curve estimated by O'Reilly and Whelan (2005):

$$\pi_t = \alpha + \rho\pi_{t-1} + \sum_{j=1}^{p} \psi_j \Delta\pi_{t-j} + \gamma x_t + \sum_{j=1}^{k} \lambda_j \Delta x_{t-j} + \delta' z_t + \varepsilon_t, \quad (1)$$

where quarterly inflation π_t (measured in annualized percentage points) is a function of its own lags (Δ denotes the first-difference operator), the output gap x_t, and a vector of supply shocks z_t. For the latter we consider quarter-on-quarter growth rates of the oil price and of the nominal effective exchange rate of the euro, denoted as o_t and e_t, respectively.[7] These shocks are included in the same way as the output gap x_t; i.e., z_t consists of contemporaneous levels and first differences up to orders l and m, such that $z_t = (o_t, \Delta o_t, \ldots, \Delta o_{t-l}, e_t, \Delta e_t, \ldots, \Delta e_{t-m})'$. Both e_t and o_t are demeaned prior to inclusion and, given that the output-gap measure has mean zero by construction, the long-run mean of inflation in (1) is given by $\alpha/(1-\rho)$. Following O'Reilly and Whelan (2005) and others, we interpret ρ as a measure of inflation persistence. Compared with O'Reilly and Whelan (2005), we allow for lags in the output-gap variable and include a number of additive price shocks. Thus, the relationship resembles the triangle model advocated by Gordon (1997).

We specify the linear Phillips curve in (1) for both the GDP deflator and HICP inflation, including quarterly dummies for the latter inflation measure in order to capture its seasonal behavior. Furthermore, we include an additive outlier dummy for 1976:Q2 to capture the spike in the GDP deflator in that quarter resulting from inflation spikes in some countries like Italy and Spain, largely associated with the consequences of the currency crises experienced in those countries over that period. For the output gap x_t, we use

[7] Some studies, including Aguiar and Martins (2005) and Rudd and Whelan (2005), have used the imported goods deflator (in its deviations from overall inflation) as a proxy for supply-side shocks. However, for the euro area such a variable is not available from 1970. The series for the import deflator that are available from 1970 (such as from the AWM database) include intra-euro-area trade, while series for the extra-euro-area import deflator are available only from the 1990s.

the first principal component of six measures of economic slack, as described in the previous section. The maximum number of lags for all variables is four, with specific lag orders chosen by combining the information from the Akaike, Schwarz (or Bayesian), and Hannan-Quinn criteria, denoted as AIC, BIC, and HQ. All models are estimated using an effective sample period from 1971:Q4 through 2005:Q4 ($T = 137$ observations).

In the process of developing a linear Phillips-curve equation for the GDP deflator and HICP inflation, it turns out that the resulting specification for the former inflation measure is considerably simpler, in the sense that the supply shocks z_t do not enhance the explanatory power of the model, while they are important for HICP inflation. For that reason we proceed by first considering a Phillips-curve specification for the GDP deflator, excluding the additive price shocks, and subsequently modeling the relationship between the GDP deflator and HICP inflation using a bridge equation, which also takes into account the additive price shocks.

The appropriate lag orders are selected by varying $p \in \{0, \ldots, 4\}$ and $k \in \{-1, 0, \ldots, 4\}$, where $p = 0$ ($k = -1$) indicates that no first differences of inflation (the output gap) are included in the model. AIC selects $p = 3$ and $k = 3$, while both BIC and HQ select $p = 3$ and $k = -1$. Upon estimating both specifications, we find that the first differences of the output gap do not add substantially to the model fit, such that we settle for the more parsimonious model, which only includes its contemporaneous level. The resulting model is given by

$$\begin{aligned}\hat{\pi}_t = \underset{(0.151)}{0.053} + \underset{(0.029)}{0.978}\ \pi_{t-1} - \underset{(0.080)}{0.493}\ \Delta\pi_{t-1} - \underset{(0.090)}{0.314}\ \Delta\pi_{t-2} \\ - \underset{(0.100)}{0.370}\ \Delta\pi_{t-3} + \underset{(0.067)}{0.280}\ x_t,\end{aligned} \tag{2}$$

$\hat{\sigma}_\pi = 3.96$, $\hat{\sigma}_\varepsilon = 1.20$, SK $= 0.38$, EK $= 1.43$, LJB $= 15.1(5.0 \times 10^{-4})$, ARCH(1) $= 0.23(0.63)$, ARCH(4) $= 7.12(0.13)$, $\mathrm{LM_{SI}}(1) = 0.29(0.59)$, $\mathrm{LM_{SI}}(4) = 1.47(0.22)$, AIC $= 0.483$, BIC $= 0.654$,

where heteroskedasticity-consistent standard errors are given in parentheses below the parameter estimates; $\hat{\sigma}_\pi$ is the standard deviation of the dependent variable; $\hat{\sigma}_\varepsilon$ is the residual standard deviation; SK and EK are residual skewness and excess kurtosis, respectively; LJB is the Lomnicki-Jarque-Bera test of normality of the residuals; ARCH(q) is the LM test of no ARCH effects up to order q in the residuals; and $\text{LM}_{\text{SI}}(m)$ is the Breusch-Godfrey test for no residual autocorrelation up to and including lag m. The numbers in parentheses following the test statistics are p-values.

The linear model seems adequate in that the errors are serially uncorrelated and homoskedastic, whereas the skewness and excess kurtosis are caused entirely by large residuals in 1973:Q1 and 1992:Q1. From this linear specification, inflation appears to be highly persistent with $\hat{\rho} = 0.978$. The coefficient of the output-gap level has the expected positive sign with $\hat{\gamma} = 0.280$.

5. Instability and Nonlinearity

In this section we assess the stability and linearity of the Phillips-curve specification for the GDP deflator discussed above. A relevant issue in this analysis is that nonlinearity and time-varying parameters generally are difficult to distinguish. In addition, instability in one part of the model may spuriously suggest instability in other parts as well. For example, a structural change in the mean of inflation, when neglected, may give the impression that inflation persistence has changed. In sum, analyzing the linearity and stability of the Phillips curve requires a well-structured and comprehensive approach. For that purpose, we adopt the methodology underlying the time-varying smooth transition (TV-STR) models as developed in Lundbergh, Teräsvirta, and van Dijk (2003). TV-STR models allow for nonlinearity and time-varying parameters simultaneously, while a modeling procedure is available for arriving at the most appropriate empirical specification; see also van Dijk, Teräsvirta, and Franses (2002) for a detailed discussion. This involves the application of a battery of diagnostic tests to a given model specification, including tests for nonlinearity and time-varying parameters, and expanding the model in the direction for which the statistical evidence is most convincing.

We start from the linear specification for the GDP deflator as given in (2). Among other misspecification tests, we separately test the stability and linearity of the intercept α, the persistence parameter ρ, and the output-gap coefficient γ as follows.

Stability of a given coefficient θ for a given variable v_t in the model is tested against the alternative of a single, gradual structural change of the form

$$\theta_t = \theta_1(1 - G(t;\xi,\tau)) + \theta_2 G(t;\xi,\tau), \tag{3}$$

where $G(t;\xi,\tau)$ is the logistic function

$$G(t;\xi,\tau) = \frac{1}{1+\exp(-\xi(t-\tau))}, \qquad \xi > 0, \tag{4}$$

which changes monotonically from 0 to 1 as t increases such that θ_t changes from θ_1 to θ_2. The restriction on the parameter ξ, which governs the smoothness of the parameter change, is for identification purposes only. The parameter τ determines the location of the shift in θ_t, in the sense that $G(t;\xi,\tau) = 0.5$ when $t = \tau$. The null hypothesis of stability can be formulated as either $\xi = 0$ or $\theta_1 = \theta_2$. In both cases, the testing problem is nonstandard due to the presence of unidentified nuisance parameters under the null hypothesis. This can be remedied by approximating the logistic function $G(t;\xi,\tau)$ by means of a low-order Taylor approximation around the point $\xi = 0$, giving rise to an auxiliary regression including terms $v_t t$, $v_t t^2$, $v_t t^3, \ldots$. This can be estimated using least squares, and a standard F-test for the joint significance of the coefficients of the auxiliary regressors provides a test for stability.

Linearity of the relationship between π_t and v_t is tested against the same alternative (3), except that in the logistic function $G(\cdot)$ in (4), time t is replaced by another observable variable s_t, which then governs the switching of θ_t between its two extreme values, θ_1 and θ_2. Here we consider nonlinear specifications with the first lag of the level and first difference of inflation, and the current level and change of the slack measure as transition variables; i.e., $s_t \in \{\pi_{t-1}, \Delta\pi_{t-1}, x_t, \Delta x_t\}$. More details about the diagnostic tests for time-varying parameters and nonlinearity can be found in Eitrheim and Teräsvirta (1996). Medeiros and Veiga (2003) develop analogous test statistics for examining the constancy and linearity of the residual variance σ_ε^2, which we also employ here.

Table 2 reports p-values of the diagnostic tests of stability and linearity applied to the different components in the linear specification for the GDP deflator. We observe that several null hypotheses are rejected—in particular, stability of the intercept α, the persistence parameter ρ, and the slope of the curve γ. The evidence for structural change in the conditional variance σ_ε^2 is less convincing. All three types of possible structural change signaled by the diagnostic tests seem plausible and have been documented in previous literature; see section 2. Given that the p-value of the stability tests for α are smallest, we proceed with estimating a model that incorporates a change in the intercept, thereby allowing for a shift in the long-term mean of inflation. This appears plausible given the substantial changes in monetary policy regimes and, in particular, in the level of inflation targets experienced by euro-area countries over the course of the past three decades. The specification of the model thus is as follows:

$$\pi_t = \alpha_t + \rho\pi_{t-1} + \sum_{j=1}^{p} \psi_j \Delta\pi_{t-j} + \gamma x_t + \sum_{j=1}^{k} \lambda_j \Delta x_{t-j} + \varepsilon_t, \tag{5}$$

where α_t is now time varying according to (3); i.e.,

$$\alpha_t = \alpha_1(1 - G(t;\xi,\tau)) + \alpha_2 G(t;\xi,\tau), \tag{6}$$

with $G(t;\xi,\tau)$ given by (4), such that $\alpha_1/(1-\rho)$ and $\alpha_2/(1-\rho)$ are the long-run means of inflation before and after the change, respectively, and can be interpreted as the central bank inflation targets during those periods. The lag orders p and k are, once again, selected on the basis of the Akaike, Schwarz, and Hannan-Quinn information criteria, all of which indicate that $p = 3$ and $k = -1$ is the preferred specification. The model is estimated with nonlinear least squares, which yields the following results:

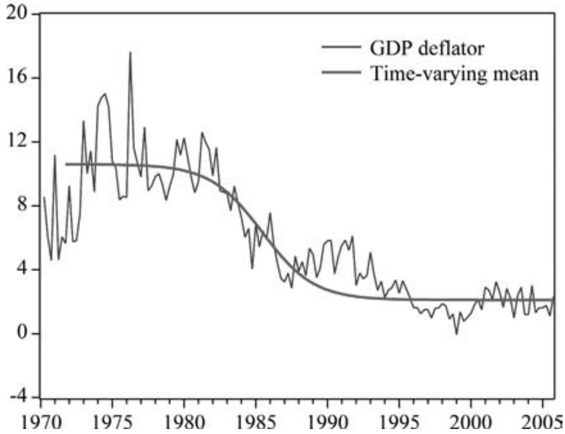
$$\begin{aligned}\hat{\pi}_t = {} & \underset{(0.950)}{3.247}\,(1 - G(t;\hat{\xi},\hat{\tau})) + \underset{(0.219)}{0.643}\,(G(t;\hat{\xi},\hat{\tau}) + \underset{(0.086)}{0.694}\,\pi_{t-1} \\ & - \underset{(0.101)}{0.330}\,\Delta\pi_{t-1} - \underset{(0.097)}{0.176}\,\Delta\pi_{t-2} - \underset{(0.085)}{0.291}\,\Delta\pi_{t-3} + \underset{(0.065)}{0.275}\,x_t,\end{aligned} \tag{7}$$

Table 2. LM-Type Tests for Nonlinearity and Time-Varying Parameters in Phillips-Curve Specifications for GDP Deflator

Transition Variable s_t	$k =$	Linear Model 1	Linear Model 2	Linear Model 3	Model with Change in Mean 1	Model with Change in Mean 2	Model with Change in Mean 3	Model with Change in Mean and Slope 1	Model with Change in Mean and Slope 2	Model with Change in Mean and Slope 3
Intercept α										
π_{t-1}		—	0.599	0.790	—	0.200	0.174	—	0.242	0.163
$\Delta\pi_{t-1}$		—	0.953	0.772	—	0.431	0.408	—	0.194	0.355
x_t		—	0.355	0.190	—	0.998	0.823	—	0.385	0.245
Δx_t		0.266	0.433	0.299	0.510	0.667	0.295	0.688	0.667	0.239
t		0.000	0.000	0.000	0.860	0.928	0.607	0.982	0.909	0.983
Persistence ρ										
π_{t-1}		0.175	0.307	0.026	0.173	0.063	0.097	0.133	0.072	0.064
$\Delta\pi_{t-1}$		0.586	0.790	0.162	0.417	0.239	0.176	0.396	0.386	0.179
x_t		0.008	0.008	0.019	0.052	0.095	0.196	0.369	0.156	0.083
Δx_t		0.289	0.445	0.336	0.462	0.704	0.630	0.195	0.366	0.444
t		0.001	0.001	0.003	0.525	0.393	0.602	0.734	0.548	0.731
Slope γ										
π_{t-1}		0.007	0.023	0.057	0.047	0.141	0.167	0.625	0.051	0.032
$\Delta\pi_{t-1}$		0.008	0.028	0.067	0.052	0.149	0.217	0.560	0.090	0.090
x_t		0.355	0.190	0.340	0.998	0.823	0.823	0.385	0.245	0.459
Δx_t		0.851	0.091	0.110	0.903	0.468	0.437	0.748	0.662	0.238
t		0.001	0.002	0.001	0.004	0.011	0.028	0.993	0.994	0.914
Residual Variance σ_ε^2										
π_{t-1}		0.100	0.103	0.160	0.047	0.026	0.025	0.142	0.115	0.117
$\Delta\pi_{t-1}$		0.119	0.171	0.111	0.049	0.018	0.025	0.147	0.115	0.117
x_t		0.213	0.164	0.304	0.466	0.302	0.304	0.249	0.300	0.176
Δx_t		0.964	0.026	0.038	0.368	0.036	0.060	0.588	0.137	0.244
t		0.106	0.127	0.080	0.067	0.074	0.027	0.178	0.164	0.119

Note: The table presents p-values of F-tests for (remaining) nonlinearity and instability in Phillips-curve specifications for quarterly inflation based on the euro-area GDP deflator for the period 1971:Q4–2005:Q4. The headings "Linear Model," "Model with Change in Mean," and "Model with Change in Mean and Slope" refer to the specifications in (2), (7), and (11), respectively. Tests are conducted for the intercept α (first panel), the persistence parameter ρ (second panel), the slope coefficient γ (third panel), and the residual variance σ_ε^2 (fourth panel). Tests are based on auxiliary regressions involving terms $v_t s_t$, $v_t s_t^2, \ldots, v_t s_t^k$, where v_t is a constant, lagged inflation π_{t-1} or the output gap x_t and s_t is the transition variable in the logistic function (4) under the alternative.

Figure 5. GDP Deflator and Time-Varying Mean



Note: The graph shows annualized quarter-on-quarter inflation rates for the euro-area GDP deflator and the time-varying mean in the Phillips-curve specification (7).

with

$$G(t;\hat{\xi},\hat{\tau}) = (1+\exp(\underset{(0.061)}{-0.138}(t-\underset{(3.72)}{53.8})))^{-1}, \tag{8}$$

$\hat{\sigma}_\pi = 3.96$, $\hat{\sigma}_\varepsilon = 1.12$, SK $= 0.34$, EK $= 0.55$, LJB $= 4.34(0.11)$, ARCH(1) $= 0.55(0.46)$, ARCH(4) $= 14.9(0.01)$, $\mathrm{LM_{SI}}(1) = 0.12(0.73)$, $\mathrm{LM_{SI}}(4) = 1.62(0.17)$, AIC $= 0.344$, BIC $= 0.514$.

The reduction in the intercept α_t is large from $\alpha_1 = 3.247$ to $\alpha_2 = 0.643$, implying a decline in the long-run mean of annualized inflation from 10.6 percent before the change to 2.1 percent thereafter. The time-varying inflation mean is plotted in figure 5, showing that the decline occurred rather gradually during the 1980s. This is broadly in line with existing literature, which dates the Great Disinflation in the early 1980s; see Cecchetti et al. (2007), among others. The second prominent feature of this specification is that allowing for a time-varying mean substantially reduces inflation persistence. The estimate of ρ in (5) is 0.694 compared with 0.978 in the specification with constant mean in (2), implying a reduction in the half-life of shocks to inflation from thirty-one to just two quarters. Finally, note that the estimated coefficient of the slack measure, $\hat{\gamma} = 0.275$, is essentially unchanged compared with the linear specification.

Table 2 reports diagnostic tests for the model with time-varying inflation mean, including tests for remaining nonlinearity and time-varying parameters. Several interesting results emerge. First, the previous evidence for time variation in inflation persistence has disappeared completely, which is in line with results of the IPN (see Altissimo, Ehrmann, and Smets 2006). Second, the single monotonic change in the intercept appears sufficient to capture the changes in the mean of inflation, as we find no statistical evidence for additional instability in the intercept. This result is somewhat surprising, as figure 5 suggests that after the large decline during the 1980s, inflation increased again during a short period around 1990, which was followed by a further downward shift to the current level of around 2 percent due to the implementation of the Maastricht Treaty and the convergence toward EMU. Third, the null hypothesis of stability of the slope parameter γ continues to be strongly rejected. This is in line with theoretical priors indicating a possible link between the level of inflation and the frequency of price adjustment, which affects the slope of the Phillips curve (Dotsey, King, and Wolman 1999). Based on the results from the various diagnostic tests, we proceed with estimating the following model, which allows for a change in slope in addition to the change in intercept:

$$\pi_t = \alpha_t + \rho\pi_{t-1} + \sum_{j=1}^{p} \psi_j \Delta\pi_{t-j} + \gamma_t x_t + \sum_{j=1}^{k} \lambda_j \Delta x_{t-j} + \varepsilon_t, \quad (9)$$

where α_t evolves according to (6), and the slope coefficient γ_t is now time varying and follows

$$\gamma_t = \gamma_1(1 - G(t;\zeta,\kappa)) + \gamma_2 G(t;\zeta,\kappa). \quad (10)$$

We obtain the following estimation results for this model:

$$\begin{aligned}\hat{\pi}_t = \underset{(0.916)}{2.742}\,(1 - G(t;\hat{\xi},\hat{\tau})) + \underset{(0.207)}{0.454}\,G(t;\hat{\xi},\hat{\tau}) + \underset{(0.084)}{0.748}\pi_{t-1} \\ - \underset{(0.099)}{0.408}\,\Delta\pi_{t-1} - \underset{(0.107)}{0.259}\,\Delta\pi_{t-2} - \underset{(0.080)}{0.327}\,\Delta\pi_{t-3} \\ + [\underset{(0.109)}{0.466}\,(1 - G(t;\hat{\zeta},\hat{\kappa})) + \underset{(0.060)}{0.134}\,G(t;\hat{\zeta},\hat{\kappa})]x_t,\end{aligned} \tag{11}$$

with

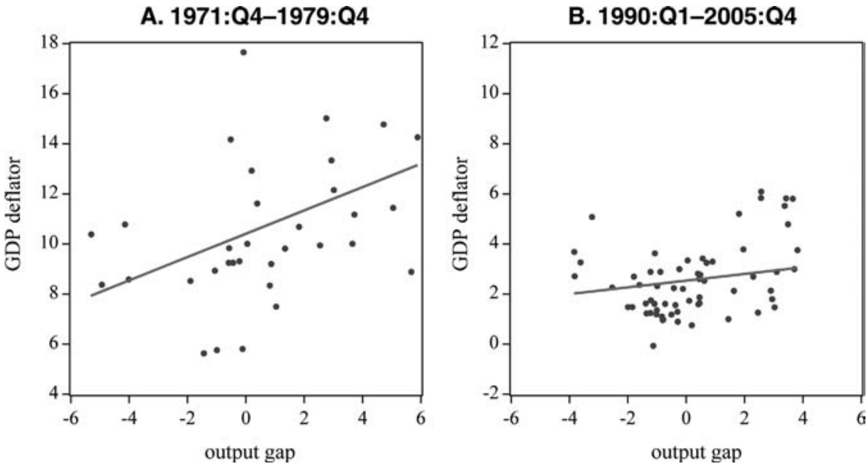
$$G(t;\hat{\xi},\hat{\tau}) = (1 + \exp(-\underset{(0.042)}{0.081}\,(t - \underset{(0.69)}{50.1})))^{-1}, \tag{12}$$

$$G(t;\hat{\zeta},\hat{\kappa}) = (1 + \exp(-\underset{(0.048)}{4.20}\,(t - \underset{(0.050)}{32.6})))^{-1}, \tag{13}$$

$\hat{\sigma}_\pi = 3.96$, $\hat{\sigma}_\varepsilon = 1.07$, SK $= 0.19$, EK $= 0.79$, LJB $= 4.44(0.11)$, ARCH(1) $= 0.39(0.53)$, ARCH(4) $= 13.6(0.01)$, $\text{LM}_{\text{SI}}(1) = 0.20(0.65)$, $\text{LM}_{\text{SI}}(4) = 6.37(0.17)$, AIC $= 0.272$, BIC $= 0.464$.

Two features of the model are striking. First, the reduction in the output-gap coefficient is substantial, with the slope after the break being approximately one-third of the slope before the break ($\hat{\gamma}_2 = 0.134$ compared with $\hat{\gamma}_1 = 0.466$). Second, the change in slope occurs rather abruptly, as indicated by the large estimate of ζ, and in 1979:Q4, prior to the change in the mean of inflation. Note that the timing and speed of the change in the intercept α_t are comparable to the estimates found before in (7), as shown in figure 5. The restriction that the timing and speed of the transitions of the intercept and slope are in fact identical is convincingly rejected on the basis of a likelihood ratio test. We believe that both the shift in the constant term and in the slope of the equation are related to the *change in monetary policy regime* in the first part of the 1980s and in particular to the transition from a regime of high and volatile inflation to a regime of low and stable inflation. One might conjecture that the transition in the frequency of price adjustment may have taken place as soon as the shift in monetary policy regime, already evident in 1980, was introduced, well before actual inflation started to fall

Figure 6. GDP Deflator and the Output Gap



Note: The graphs show scatter plots of the quarterly output gap against the annualized quarter-on-quarter inflation rate for the euro-area GDP deflator for the periods 1971:Q4–1979:Q4 and 1990:Q1–2005:Q4. The solid line shows the relationship $\hat{\pi}_t = \bar{\pi} + \hat{\gamma}_i x_t$ for $i = 1, 2$ where $\hat{\gamma}_i$ are the nonlinear least-squares estimates from (11) and $\bar{\pi}$ is the sample mean inflation rate over the respective subperiods.

and converge to lower levels as from the mid-1980s.[8] As a result of the shift in the slope, the Phillips curve has become significantly flatter from 1990 onward (i.e., after the adjustment of both the persistence and slope levels appears to be completed) compared with the 1970s (i.e., before these changes took place), as suggested by figure 6.

Table 2 reports diagnostic tests for remaining instability and nonlinearity for this model with time-varying intercept and slope. For most tests, the p-values are well above conventional significance levels. As regards nonlinearity, we find statistically insignificant test statistics when the transition variable is the level or first difference of the output gap; i.e., $s_t = x_t$ or Δx_t. Note that a nonlinear relationship between the inflation rate π_t and the output gap x_t with

[8]This may imply that a model in which the change in the slope of the curve is driven by the level of (trend) inflation, as in De Veirman (2007) for Japan, may not be very appropriate in the euro area.

x_t itself as the transition variable would correspond to a concave or convex functional form of the Phillips curve. Hence, we find no evidence of these commonly studied types of nonlinearity for the euro area. However, we do find some indications for the presence of nonlinearity in the relationship between inflation and the output gap, as the p-values of the linearity tests with $s_t = \pi_{t-1}$ are below 10 percent. We attempted to estimate a smooth transition regression model accordingly, but this did not give meaningful results. Hence, we accept the specification in (11) as an adequate representation of the Phillips-curve dynamics over the period 1970–2005.

5.1 Sensitivity Analysis

We perform two types of sensitivity analysis to examine the robustness of our results. First, we include the price shocks z_t in the Phillips-curve specification as in (1). The information criteria suggest to include only the contemporaneous level of the oil price shock o_t and the contemporaneous level and one lagged first difference of the exchange rate shock e_t. As already noted in section 4, we find very little role for these additive price shocks in the equation for the GDP deflator and, not surprisingly, the main results concerning the changes in mean and slope of the Phillips curve remain practically unchanged.[9]

Second, we reestimate the model in (11) by substituting each of the six individual measures of the output gap discussed in section 3 as well as their arithmetic average for the summary measure based on the first principal component used before. Table 3 presents estimates of the parameters determining the time-varying slope γ_t as defined in (10) for the different choices of the output-gap measure x_t. To account for the different amplitude of the slack measures, we report scaled coefficients $\gamma_i^* = \gamma_i \times \sigma_x$, $i = 1, 2$, where σ_x denotes the sample standard deviation of x_t. The table shows that the coefficient estimates for the principal component and the arithmetic average are very close, both for the timing and speed of the structural change of the slope coefficient as well as its magnitude before and after the change. The same holds for the three gap measures

[9] Results are not reported for brevity but are available from the authors upon request.

Table 3. Output-Gap Measures—Sensitivity Analysis

	PC	AVG	CC	PIC	BIV	BK	HP	UC
γ_1^*	1.042	1.115	1.384	1.495	1.546	0.855	0.830	0.833
γ_2^*	0.300	0.304	0.311	0.179	0.369	0.222	0.154	0.102
ζ	4.200	3.750	4.200	20.00	0.207	20.00	0.526	20.00
κ	39.60	39.60	39.30	42.80	34.00	80.40	79.60	40.10

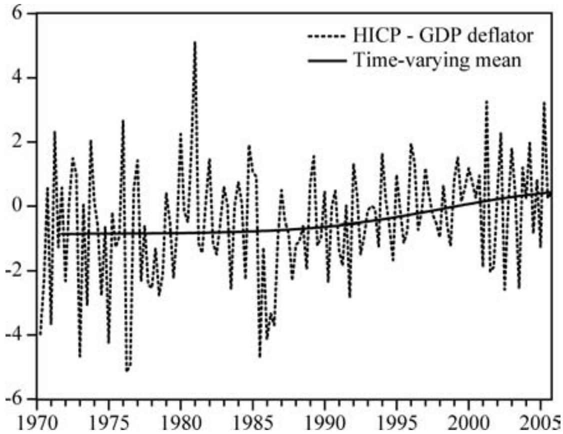
Note: The table presents estimates of the parameters in the time-varying slope γ_t defined in (10), which is used in the Phillips-curve specification given in (9) for different choices of the output-gap measure x_t. $\gamma_i^* = \gamma_i \times \sigma_x$, $i = 1, 2$, where σ_x denotes the sample standard deviation of x_t. CC, PIC, and BIV are obtained from the common cycles, the pseudo-integrated cycles, and the bivariate versions, respectively, of the multivariate unobserved-components model of Proietti, Musso, and Westermann (2007). BK denotes the Baxter-King band-pass filter, HP the Hodrick-Prescott filter, and UC a univariate unobserved-components model applied to quarterly real GDP. PC denotes the first principal component of these six measures, while AVG denotes their simple average.

based on the multivariate unobserved-components model of Proietti, Musso, and Westermann (2007). Larger differences are observed for the univariate measures based on statistical filters applied to real GDP. In particular, when using the Baxter-King band-pass filter or the Hodrick-Prescott filter, the timing of the change is dated a full decade later compared with the univariate unobserved-components model or any of the other gap measures ($\hat{\kappa} \approx 80$ as opposed to 40).

6. Modeling HICP Inflation

Although the rate of inflation derived from the GDP deflator is of great interest, the ECB's monetary policy objective of price stability is defined in terms of HICP inflation. For that reason, in this section we develop a model for HICP inflation, linking it to the GDP deflator using a so-called bridge equation, which has the difference between the HICP and GDP deflator inflation measures as the dependent variable. As discussed in sections 3 and 4, HICP inflation moves closely together with the GDP deflator inflation, but with two important differences. First, while GDP deflator inflation appears not to be affected by our measures of additive price shocks, HICP inflation

Figure 7. Difference between HICP and GDP Deflator Inflation



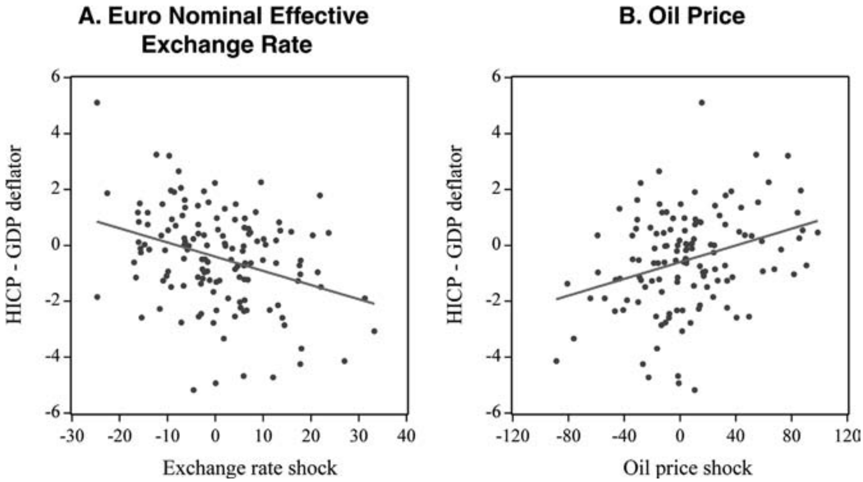
Note: The graph shows the difference between the annualized quarter-on-quarter inflation rates for the euro-area HICP and GDP deflator for the period 1970:Q2–2005:Q4. The horizontal line is the time-varying mean in the specification (14).

is. Second, while the GDP deflator is seasonally adjusted, the HICP is not.

Figure 7 plots the difference between the annualized quarter-on-quarter rate of change in the HICP and in the GDP deflator, denoted d_t. The difference between the two inflation measures appears to be stationary over the sample period that we cover, although an increase in the level seems to have occurred since 1990, approximately. Some seasonality also appears to be present in the series. The effects of price shocks become clear from figure 8, which shows scatter plots of the quarterly changes in the euro nominal effective exchange rate and in the oil price against d_t. As expected, the inflation differential is negatively related to exchange rate shocks and positively related to oil price shocks.

An appropriate model for d_t is developed using the same procedure applied to the Phillips-curve specification for the GDP deflator as discussed in sections 4 and 5. That is, we start with a linear specification of the form (1), but for d_t instead of π_t, and not including the

Figure 8. Price Shocks: Euro Nominal Effective Exchange Rate and Oil Price



Note: The graphs show scatter plots of the quarterly change in the euro nominal effective exchange rate and oil price against the difference between the annualized quarter-on-quarter inflation rates for the euro-area HICP and GDP deflator for the period 1970:Q2–2005:Q4. The solid line shows the least-squares fit of the inflation differential on a constant and the price shock.

terms involving the output gap x_t.[10] We do include the price shocks $z_t = (o_t, \Delta o_t, \ldots, \Delta o_{t-l}, e_t, \Delta e_t, \ldots, \Delta e_{t-m})'$ and, in addition, a set of centered seasonal dummies $D_t = (D^*_{1,t}, D^*_{2,t}, D^*_{3,t})' \equiv (D_{1,t} - D_{4,t}, D_{2,t} - D_{4,t}, D_{3,t} - D_{4,t})'$, where $D_{s,t}$, $s = 1, \ldots, 4$ are quarterly dummy variables, with $D_{s,t} = 1$ when time t corresponds with quarter s and $D_{s,t} = 0$ otherwise. Finally, additive outlier dummies are included for 1976:Q2 as before, as well as for 1974:Q1 to handle the extremely large oil price shocks that occurred at that time.

[10]By definition, the difference between HICP and GDP inflation is the inflation rate on imports weighted by their share in the consumers' basket. As Galí and Gertler (1999) point out, the output gap x_t acts as a proxy for marginal costs in the determination of prices by monopolistically competitive firms producing value added (GDP), whereas imported inflation is taken as given by these firms. Hence, x_t does not necessarily cancel in the differential inflation d_t. We did attempt to include the output-gap measure in the model for d_t, but it turned out to be insignificant.

Based on the information criteria, the lag orders are set equal to $p = 4$, $l = 0$, and $m = -1$; i.e., we include the contemporaneous level and first difference of the oil price shock o_t and only the contemporaneous level of the exchange rate shock e_t. The estimated model is subjected to the usual misspecification tests for nonlinearity and parameter instability.[11] The test results indicate instability in the intercept of the model, reflecting the change in level of d_t, as well as instability in the coefficients of the quarterly dummies D_t, suggesting that the seasonal pattern also may have changed. No signs for instability or nonlinearity in the effects of the shocks o_t and e_t are found. (Sequentially) Incorporating the change in intercept and seasonality into the model, we finally arrive at the following estimated model:

$$\begin{aligned}
\hat{d}_t = {} & \underset{(0.317)}{-0.812}(1 - G(t;\hat{\xi},\hat{\tau})) + \underset{(1.397)}{0.637}\, G(t;\hat{\xi},\hat{\tau}) + \underset{(0.158)}{0.047}\, d_{t-1} \\
& - \underset{(0.143)}{0.075}\, \Delta d_{t-1} - \underset{(0.129)}{0.029}\, \Delta d_{t-2} + \underset{(0.117)}{0.074}\, \Delta d_{t-3} + \underset{(0.082)}{0.224}\, \Delta d_{t-4} \\
& + \underset{(0.0012)}{0.0045}\, o_t + \underset{(1.5 \times 10^{-4})}{9.4 \times 10^{-5}}\, \Delta o_t - \underset{(0.011)}{0.042}\, e_t \\
& + \left[\underset{(1.268)}{-2.362}\, D^*_{1,t} - \underset{(0.866)}{1.019}\, D^*_{2,t} + \underset{(0.835)}{2.412}\, D^*_{3,t}\right](1 - G(t;\hat{\zeta},\hat{\kappa})) \\
& + \left[\underset{(0.262)}{0.679}\, D^*_{1,t} + \underset{(0.229)}{0.499}\, D^*_{2,t} - \underset{(0.193)}{0.868}\, D^*_{3,t}\right] G(t;\hat{\zeta},\hat{\kappa}),
\end{aligned} \tag{14}$$

with

$$G(t;\hat{\xi},\hat{\tau}) = (1 + \exp(\underset{(0.075)}{-0.055}\,(t - \underset{(1.99)}{107.6})))^{-1}, \tag{15}$$

$$G(t;\hat{\zeta},\hat{\kappa}) = (1 + \exp(\underset{(0.039)}{-20.00}(t - \underset{(0.005)}{6.9})))^{-1}, \tag{16}$$

$\hat{\sigma}_d = 1.73$, $\hat{\sigma}_\varepsilon = 1.17$, SK $= -0.19$, EK $= 0.35$, LJB $=$ 1.49(0.47), ARCH(1) $=$ 0.72(0.40), ARCH(4) $=$ 5.16(0.27),

[11] Results are not reported for brevity but are available from the authors upon request.

$\text{LM}_{\text{SI}}(1) = 1.12(0.29)$, $\text{LM}_{\text{SI}}(4) = 0.67(0.61)$, AIC = 0.583, BIC = 0.966.

Several features of the model are noteworthy. First, the model explains more than half of the variation in the inflation differential and appears adequate, as the usual diagnostic tests do not indicate any obvious misspecification. Second, the change in mean occurs gradually and is centered around 1997; see also figure 7. The mean inflation differential changes from −0.85 percentage points before the change to 0.67 percentage points after. The latter should be interpreted with caution, however, as the function $G(t; \hat{\xi}, \hat{\tau})$ only takes the value 0.85 at the end of our sample period such that the change is not completed. Third, the estimates of the parameters in the second logistic function $G(t; \hat{\zeta}, \hat{\kappa})$ indicate that the change in seasonality occurs rapidly during the first half of 1973. Hence, the instability in the seasonal pattern appears to be due to a few erratic observations early in the sample period. Fourth, the oil price shock o_t has a significant positive effect on the inflation differential, consistent with the idea that an oil price increase leads to higher consumer prices but does not affect the GDP deflator. Similarly, the significantly negative coefficient for the exchange rate shock e_t suggests that consumer prices are influenced by changes in the euro exchange rate.

It is worth mentioning that the finding of no evidence of a *direct* asymmetric impact of oil price shocks *on inflation* is not necessarily inconsistent with an *overall* asymmetric impact once the transmission channel through the output gap is taken into account, if oil prices do have an asymmetric impact on demand conditions.

7. Conclusions

This paper has aimed at providing a comprehensive analysis of the stability and linearity of the euro-area Phillips curve, a question that is of obvious policy relevance in Europe, where a stable rate of inflation appears to coexist with a seemingly high level of spare capacity. The main results of the study are three. First, there is strong evidence, quite unsurprisingly, of a shift in the mean of euro-area inflation, with the change occurring quite gradually toward the middle of the 1980s. Second, there is also strong evidence of a shift in the slope of the curve, again occurring in the 1980s but somewhat earlier and much more abruptly. As a result of this shift, the curve becomes significantly flatter, consistent with the idea that the frequency of price adjustment is negatively related to the mean of inflation. Third, once we correct for this time variation in the parameters, we find no significant evidence of nonlinearity in the curve—in particular, in relation to the output gap. Hence, we conclude that the Phillips "curve" is, at least in the euro area, indeed a "line." The main policy implication of our study is, therefore, that there is at least no convincing evidence of the existence of a "free lunch" for monetary policy, whereby the central bank is able to stimulate economic activity without creating inflationary pressure.

Further analysis at the level of the individual countries in the euro area could be useful in order to ascertain whether there is any interesting heterogeneity in the stability and functional form of the Phillips curve. In particular, it appears interesting to compare low-inflation (e.g., Germany) and high-inflation (e.g., Italy) countries over a longer sample period, before the start of the monetary union. This appears to be an interesting avenue for future research.

References

Aguiar, A., and M. M. F. Martins. 2005. "Testing the Significance and the Non-Linearity of the Phillips Trade-Off in the Euro Area." *Empirical Economics* 30 (3): 665–91.

Akerlof, G. A., W. T. Dickens, and G. L. Perry. 1996. "The Macroeconomics of Low Inflation." *Brookings Papers on Economic Activity* 1: 1–59.

Altissimo, F., M. Ehrmann, and F. Smets. 2006. "Inflation Persistence and Price-Setting Behaviour in the Euro Area—A Summary of the IPN Evidence." ECB Occasional Paper No. 46.

Angeloni, I., L. Aucremanne, and M. Ciccarelli. 2006. "Price Setting and Inflation Persistence: Did EMU Matter?" *Economic Policy* 21 (46): 353–87.

Baghli, M., C. Cahn, and H. Fraisse. 2007. "Is the Inflation-Output Nexus Asymmetric in the Euro Area?" *Economics Letters* 94 (1): 1–6.

Barnes, M. L., and G. P. Olivei. 2003. "Inside and Outside Bounds: Threshold Estimates of the Phillips Curve." *New England Economic Review* (Federal Reserve Bank of Boston) (2003): 3–18.

Baxter, M., and R. G. King. 1999. "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series." *Review of Economics and Statistics* 81 (4): 575–93.

Blanchard, O., and J. Galí. 2007. "The Macroeconomic Effects of Oil Price Shocks: Why Are the 2000s So Different from the 1970s?" NBER Working Paper No. 13368 (September).

Borio, C. E. V., and A. J. Filardo. 2007. "Globalisation and Inflation: New Cross-Country Evidence on the Global Determinants of Domestic Inflation." BIS Working Paper No. 227.

Camba-Méndez, G., and D. Rodriguez-Palenzuela. 2003. "Assessment Criteria for Output Gap Estimates." *Economic Modelling* 20 (3): 528–61.

Cecchetti, S. G., and G. Debelle. 2006. "Has the Inflation Process Changed?" *Economic Policy* 21 (46): 311–52.

Cecchetti, S. G., P. Hooper, B. C. Kasman, K. L. Schoenholtz, and M. W. Watson. 2007. "Understanding the Evolving Inflation Process." Report presented at the U.S. Monetary Policy Forum, Washington, DC, March 9.

Clark, P. B., and D. Laxton. 1997. "Phillips Curves, Phillips Lines, and the Unemployment Costs of Overheating." IMF Working Paper No. 97/17.

Corvoisier, S., and B. Mojon. 2005. "Breaks in the Mean of Inflation: How They Happen and What To Do with Them." ECB Working Paper No. 451.

Debelle, G., and D. Laxton. 1997. "Is the Phillips Curve Really a Curve? Some Evidence for Canada, the United Kingdom, and the United States." *IMF Staff Papers* 44 (2): 249–82.

De Gregorio, J., O. Landerretche, and C. Neilson. 2007. "Another Pass-Through Bites the Dust? Oil Prices and Inflation." Central Bank of Chile Working Paper No. 417.

De Veirman, E. 2007. "Which Nonlinearity in the Phillips Curve? The Absence of Accelerating Deflation in Japan." Working Paper.

Dolado, J. J., R. Maria-Dolores, and M. Naveira. 2005. "Are Monetary-Policy Reaction Functions Asymmetric? The Role of

Nonlinearity in the Phillips Curve." *European Economic Review* 49 (2): 485–503.

Dotsey, M., R. G. King, and A. L. Wolman. 1999. "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output." *Quarterly Journal of Economics* 114 (2): 655–90.

Eitrheim, Ø., and T. Teräsvirta. 1996. "Testing the Adequacy of Smooth Transition Autoregressive Models." *Journal of Econometrics* 74 (1): 59–76.

Fabiani, S., M. Druant, I. Hernando, C. Kwapil, 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.

Fabiani, S., and R. Mestre. 2004. "A System Approach for Measuring the Euro Area NAIRU." *Empirical Economics* 29 (2): 311–41.

Fagan, G., J. Henry, and R. Mestre. 2005. "An Area-Wide Model for the Euro Area." *Economic Modelling* 22 (1): 39–59.

Filardo, A. J. 1998. "New Evidence on the Output Cost of Fighting Inflation." *Economic Review* (Federal Reserve Bank of Kansas City) (Third Quarter): 33–61.

Galí, J., and M. Gertler. 1999. "Inflation Dynamics: A Structural Econometric Analysis." *Journal of Monetary Economics* 44 (2): 195–222.

Gordon, R. J. 1997. "The Time-Varying NAIRU and Its Implications for Economic Policy." *Journal of Economic Perspectives* 11 (1): 11–32.

Ihrig, J., S. B. Kamin, D. Lindner, and J. Marquez. 2007. "Some Simple Tests of the Globalization and Inflation Hypothesis." International Finance Discussion Paper No. 891, Board of Governors of the Federal Reserve System.

Ihrig, J., M. Marazzi, and A. D. Rothenberg. 2006. "Exchange-Rate Pass-Through in the G-7 Countries." International Finance Discussion Paper No. 851, Board of Governors of the Federal Reserve System.

Laxton, D., D. Rose, and D. N. Tambakis. 1999. "The U.S. Phillips Curve: The Case for Asymmetry." *Journal of Economic Dynamics and Control* 23 (9–10): 1459–85.

Lundbergh, S., T. Teräsvirta, and D. van Dijk. 2003. "Time-Varying Smooth Transition Autoregressive Models." *Journal of Business and Economic Statistics* 21 (1): 104–21.

Medeiros, M. C., and A. Veiga. 2003. "Diagnostic Checking in a Flexible Nonlinear Time Series Model." *Journal of Time Series Analysis* 24 (4): 461–82.

Melick, W., and G. Galati. 2006. "The Evolving Inflation Process: An Overview." BIS Working Paper No. 196.

O'Reilly, G., and K. Whelan. 2005. "Has Euro-Area Inflation Persistence Changed Over Time?" *Review of Economics and Statistics* 87 (4): 709–20.

Orphanides, A., and S. van Norden. 2005. "The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time." *Journal of Money, Credit, and Banking* 37 (3): 583–601.

Proietti, T., A. Musso, and T. Westermann. 2007. "Estimating Potential Output and the Output Gap for the Euro Area: A Model-Based Production Function Approach." *Empirical Economics* 33 (1): 85–113.

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

Rudd, J., and K. Whelan. 2005. "New Tests of the New-Keynesian Phillips Curve." *Journal of Monetary Economics* 52 (6): 1167–81.

———. 2007. "Modeling Inflation Dynamics: A Critical Review of Recent Research." *Journal of Money, Credit, and Banking* 39 (s1): 155–70.

Rünstler, G. 2002. "The Information Content of Real-Time Output Gap Estimates: An Application to the Euro Area." ECB Working Paper No. 182.

Stiglitz, J. 1997. "Reflections on the Natural Rate Hypothesis." *Journal of Economic Perspectives* 11 (1): 3–10.

Stock, J. H., and M. W. Watson. 2007. "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit, and Banking* 39 (s1): 3–33.

van Dijk, D. J., D. R. Osborn, and M. Sensier. 2002. "Changes in Variability of the Business Cycle in the G7 Countries." Econometric Institute Research Report No. 2002-28, Erasmus University Rotterdam.

van Dijk, D., T. Teräsvirta, and P. H. Franses. 2002. "Smooth Transition Autoregressive Models—A Survey of Recent Developments." *Econometric Reviews* 21 (1): 1–47.