

THE OUTPUT COMPOSITION PUZZLE: A DIFFERENCE IN THE MONETARY TRANSMISSION MECHANISM IN THE EURO AREA AND THE UNITED STATES

Ignazio Angeloni
European Central Bank

Anil K Kashyap
*University of Chicago, Graduate School of Business,
Federal Reserve Bank of Chicago, and NBER*

Benoît Mojon
European Central Bank

Daniele Terlizzese
Banca d'Italia

A consensus has emerged during the last twenty years, over the way that the actions of central banks affect the economy (the monetary transmission mechanism). In a nutshell, changes in monetary policy have a persistent, though not permanent, effect on output,

Kashyap thanks the ECB, the University of Chicago George J. Stigler Center for the Study of the Economy and the State, and the National Science Foundation (through a grant administered by the National Bureau of Economic Research) for financial support. The views expressed in this paper are not necessarily shared by the institutions to which the authors are affiliated. We thank M. Aglietta, F. Altissimo, J.P. Benassy, A. Blinder, L. Christiano, M. Eichenbaum, C. Erceg, C. Evans, A. Gerali, L. Guiso, J. Pisani-Ferry, R. McCulloch, J.P. Pollin, D. Reifschneider, P. Rossi, C. Sims, F. Smets, L. Sessa and R. Wouters for helpful conversations; our discussants J. Lindé and R. Clarida and seminar participants at the Central Bank of Chile, the ECB, the University of Chicago, and the Wharton School for useful comments; A.M. Agresti, A. Gerali, F. Brayton, C. Erceg, J. Roush, F. Smets and R. Wouters for helping us in various ways to assemble some of the evidence and the calculations presented. All errors and shortcomings are our responsibility alone.

First published in the *Journal of Money, Credit, and Banking*, "Recent Developments in Monetary Macroeconomics" Conference Proceedings, November 2002, sponsored by the Federal Reserve Bank of Cleveland. Reprinted with permission.

Banking Market Structure and Monetary Policy, edited by Luis Antonio Ahumada and J. Rodrigo Fuentes, Santiago, Chile. ©2004 Central Bank of Chile.

with the output change being hump shaped; prices react with some delay, and eventually settle down to a new level, with no permanent effect on inflation.

Much of this consensus is based on the examination of the U.S. experience. Yet, recently, twelve European countries embarked upon an unprecedented grand monetary experiment. A new central bank was created from scratch and the currencies of twelve sovereign nations were replaced with the euro. A natural question is whether the consensus view on the monetary transmission mechanism holds for the euro area as well.

While we expect this question will be the subject of intense research in the future, some first answers were provided by a momentous research effort involving the staffs of the European Central Bank (ECB) and of the twelve national central banks (NCBs) forming the euro area (Angeloni, Kashyap, and Mojon, 2003). Some surprising similarities were found, together with some interesting differences.

In this paper, drawing from that body of work, we first check the robustness of the similarities. These are important because, as the euro area is only about five years old, any time series analysis of the euro area transmission necessarily uses mostly data from the previous monetary policy regime. This confounds analyses based on either synthetic data of euro area aggregates or the aggregation of country-level findings. However, some of the uncertainty over the transmission mechanism may be reduced if the time series facts that can be compiled for the euro area resemble those for the United States, a long-functioning monetary union of similar size and openness as the euro area.

The bulk of our analysis focuses on an intriguing difference between the two currency areas. In particular, we call attention to one aspect of the transmission mechanism that has previously received little attention: the composition of the output adjustments that follow a change in monetary policy. Along this dimension, an interesting contrast emerges between the euro area findings and those for the United States. In the United States, changes in consumption spending appear to be a much more important component of monetary adjustment than in the euro area (where investment spending changes appear to be preeminent). We dub this difference the *output composition puzzle*.

We see the motivation for studying the composition of the output response as threefold. First, better understanding the composition effects can improve the central bank's ability to monitor the economy.

For instance, knowledge that consumption adjustments are typically dominant in the United States would suggest that consumer behavior is what needs to be watched carefully to see whether policy changes are working through the economy in the expected way. This ultimately would help determine whether the current monetary stance is appropriate or whether policy changes are called for.

A second, broader motivation is that knowledge of the composition can improve our understanding of the factors behind the monetary transmission mechanism. As discussed later on, the differences between a dominant consumption response in the United States and a dominant investment response in the euro area could be due to a variety of institutional or legal constraints, or frictions, linked for example to the structure of financial or labor markets or to differences in the levels of social insurance. Better understanding the composition seems a useful first step to uncover the relevance of these different factors. Moreover, having identified the relevant factors, one could then discuss whether structural policies—for instance, in the financial or labor markets—might be warranted to alter these institutions.

A third, and closely related, consideration is that this analysis can be informative about the stability of the transmission mechanism. By understanding which transmission channels are dominant and which are dormant, one can decide which changes to the economy merit most attention. For instance, if the consumption response in the United States is dominant, a policymaker might conclude that paying close attention to changes in the mortgage markets is more important than studying changes in the tax treatment of depreciation.

We organize the paper into three parts. We begin with a brief review of the stylized facts about the basic statistical properties of the data and on the transmission mechanisms for the United States and euro area, showing a number of similarities.

In the next section we document the output composition puzzle, arguing that it is a robust feature of the two economies that can be confirmed using a host of statistical techniques and data.

In the following section, we provide tentative interpretations and explanations for it. We first explore the puzzle in the class of tractable dynamic stochastic general equilibrium (DSGE) models that have recently been proposed as an accurate description of the monetary policy transmission (prominent examples are Christiano, Eichenbaum, and Evans, 2001, for the United States and Smets and Wouters, 2002, for the euro area). The idea is to trace the differences in output composition to differences in “deep” parameters characterizing the two

economies. We verify that these models, in their current estimated (or calibrated) version, have trouble fully accounting for the differences in the composition of output adjustments that we observe in the data.

To do this we identify the mechanisms in the model that give rise to differences in the output composition, isolating a small subset of the models' parameters that essentially governs the output composition. The differences estimated for these parameters are too small, however, and sometimes even of the wrong sign, to fully account for the differences in the output composition between the two areas. Moreover, the mechanisms identified do not appear to be very powerful. It appears that large changes in these parameters are needed to bring the models in line with our data-based estimates of the consumption contributions to output adjustment.

Whether or not these DSGE models could be modified and reestimated to overcome these problems and account for the output composition puzzle is an issue that we leave for future research. For now, they provide us with a structural (although partial) interpretation of the uncovered differences that can be subject to independent scrutiny. Most importantly, revealing that some potential mechanisms are not enough to account for the puzzle helps direct the search for other mechanisms, so far not included in these models.

We move in this direction in the final section of the paper. There, departing from the maintained assumption in the DSGE models that agents are fully insured against various shocks, we explore differences in employment and income risk to see whether the lack of these kinds of insurance might be responsible for the differences. The evidence is ambiguous but there are some hints that more complete social insurance in the euro area might play a role in resolving our puzzle. Overall, we tentatively attribute the origin of the puzzle to differences in the behavior of consumers rather than in the behavior of firms (through their investment decisions).

1. BASIC FACTS ON MONETARY TRANSMISSION IN THE UNITED STATES AND EURO AREA

A vast literature of the monetary transmission mechanism exists, with excellent recent surveys provided by the papers in the 1995 symposium in the *Journal of Economic Perspectives* (Bernanke and Gertler, 1995; Meltzer, 1995; Obstfeld and Rogoff, 1995; Taylor, 1995), Christiano, Eichenbaum, and Evans (1999), Mankiw (2001), and Bean,

Larsen, and Nikolov (2003). Rather than rehashing the evidence reviewed in these papers, we focus on whether the long U.S. expansion in the 1990s has changed anything and compare the latest U.S. results to some recent findings for the euro area. As they are relatively less known, we start by taking a look at the euro area data.

1.1 Introduction to the Euro Area Data

One major challenge in analyzing the transmission mechanism in the euro area is the data difficulties. The euro area has only had a single monetary policy for about five years. So time series analysis of macroeconomic variables during this time period is not feasible.

Combining the post-ECB data with historical data is also difficult. For one thing, many countries that now use the euro do not have full quarterly data on many relevant macroeconomic series. For example, quarterly data for inventory investment and durable consumption are simply not available for most countries. Furthermore, quarterly euro area trade figures net of trade flows within the euro area are only available from 1992 onwards. Thus, there are certain questions that cannot even be considered.

More fundamentally, it is legitimate to question whether aggregating the country data for the euro area countries prior to the adoption of the euro even makes sense. This was obviously not a single economy with a common monetary policy prior to 1999, though the transition to the single currency and the likely ensuing changes in agents' behavior were gradual. So one might prefer to analyze the member countries separately and then aggregate the findings to the euro area level.

But this approach also has problems. First, the data limitations are substantial even at the country level. Second, we are chiefly interested in how the member countries would respond to common monetary actions. Given that in the historical sample there was no common monetary policy, we need to adjust the country level results anyway (for instance, by imposing a common monetary reaction function in the analysis). Recognizing these problems, we analyze both the synthetic data for the euro area and country-level evidence.¹

1. The euro area data used in this study are taken from Fagan, Henry, and Mestre (2001). Updates of these data along with a number of other statistical data on the euro area real and financial sectors are available at the ECB website (www.ecb.int).

We begin by reporting some summary descriptive material on the euro area data. Tables 1 and 2, reproduced from Agresti and Mojon (2003), presents a set of descriptive statistics for the (detrended) euro area data along with similar statistics for the United States, which serves as a benchmark. The euro area data are only available from 1970 onwards, so for comparison purposes we show findings for both regions from this date through 2000; in later sections we take advantage of earlier U.S. data where available.

Three main features of these results stand out. First, the *absolute level* of the volatility of gross domestic product (GDP) in the euro area is lower than in the United States.² Second, if measured *relative* to GDP, the volatility of the main domestic demand components appear to be broadly similar in the two economies; of relevance for our later findings is the fact that the relative volatilities of consumption and investment are similar in both currency areas. This does not appear to be true for inflation (as measured by consumer price indices), whose volatility appears to be much lower in the euro area.

Third, the dynamic cross and auto-correlations between the main macroeconomic variables display many striking similarities across the two economies. For instance, the serial correlation properties of GDP and the price deflators, as well as the lead-lag patterns of the cross-correlations between GDP and its components, interest rates and credit aggregates are all broadly similar.

There are also several differences. The one that we find most intriguing is that stock prices appear to be strongly positively correlated with future output in the United States, contrary to what is found for the euro area. This could result from the small size of the stock market in continental Europe over most of the sample period. We do not have obvious explanations for the other dissimilarities.³

2. In this context it should be noted, however, that the volatility of U.S. GDP has declined over time. See Stock and Watson (2003) for a survey, and Ahmed, Levin, and Wilson (2002), Kahn, McConnell, and Pérez-Quirós (2002), Boivin and Giannoni (2002), Clarida, Galí, and Gertler (2000), and Ramey and Vine (2003) for competing explanations of this reduction in macroeconomic instability.

3. For instance, we do not have interpretations for the following findings: (1) that the correlation between past GDP and current inflation tends to be lower in the euro area; (2) that the sign of the correlation between current inflation and future GDP growth quickly becomes negative in the United States, while it remains positive in the euro area; (3) that M1 seems a better leading indicator of output in the euro area than in the United States; and (4) that real estate prices exhibit very different lead and lag correlations with GDP in the two economies.

Table 1. Descriptive Statistics for Cyclical Components of Euro Area Time Series, 1970–2000^a

Variable (t)	Standard deviation		Cross-correlation with GDP (t + k)										
	absolute	relative/GDP	k: -4	-3	-2	-1	0	1	2	3	4		
GDP	0.84	1.0	-0.19	0.18	0.58	0.88	1.00						
Consumption	0.55	0.7	-0.13	0.09	0.37	0.63	0.79	0.80	0.66	0.40	0.09		
Investment	1.85	2.2	0.06	0.34	0.62	0.81	0.86	0.75	0.51	0.21	-0.09		
Cumulated inventories	2.40	2.9	0.65	0.83	0.82	0.59	0.22	-0.19	-0.52	-0.70	-0.70		
GDP deflator (level)	0.58	0.7	0.29	0.27	0.16	-0.04	-0.30	-0.55	-0.72	-0.76	-0.67		
CPI (level)	0.68	0.8	0.28	0.26	0.16	-0.03	-0.26	-0.50	-0.66	-0.72	-0.66		
CPI (inflation)	0.31	0.4	0.35	0.34	0.27	0.21	0.20	0.26	0.31	0.30	0.19		
Stock prices	12.00	14.3	-0.10	-0.07	-0.01	0.05	0.08	0.06	0.01	-0.03	-0.02		
Real estate prices	1.36	1.6	0.53	0.52	0.50	0.45	0.39	0.31	0.20	0.06	-0.08		
Short-term rate nominal	1.09	1.3	0.27	0.54	0.73	0.76	0.61	0.30	-0.08	-0.43	-0.67		
Short-term rate real	0.76	0.9	0.49	0.65	0.68	0.55	0.26	-0.11	-0.43	-0.61	-0.59		
Long-term rate nominal	0.57	0.7	0.22	0.38	0.48	0.47	0.33	0.09	-0.17	-0.37	-0.46		
Yield spread	0.83	1.0	-0.20	-0.45	-0.63	-0.68	-0.58	-0.34	-0.01	0.32	0.56		
Real effective exchange rate	3.58	4.3	0.22	0.33	0.36	0.30	0.17	0.01	-0.12	-0.18	-0.18		
DM-USD exchange rate	5.23	6.2	0.13	0.36	0.56	0.61	0.48	0.22	-0.08	-0.34	-0.46		
M1	1.00	1.2	-0.22	-0.26	-0.20	-0.05	0.16	0.39	0.58	0.68	0.67		
M3	0.72	0.9	0.45	0.23	0.01	-0.17	-0.26	-0.27	-0.19	-0.06	0.07		
Total loans	0.85	1.0	0.59	0.55	0.48	0.37	0.23	0.10	0.00	-0.06	-0.08		
									Cross-correlation with own (t + k)				
CPI (level)	0.68	0.8	0.33	0.55	0.77	0.77	1.00						
GDP deflator (level)	0.31	0.4	0.27	0.50	0.74	0.74	1.00						

Source: Agresti and Mojon (2003).

a. Standard deviation of and cross-correlation between the cyclical components of macroeconomic time series. The cyclical component was obtained from a band-pass filter BPF(6,40.8) à la Baxter and King (1999), and the data sources are fully described in Agresti and Mojon (2001, appendix 1).

Table 2. Descriptive Statistics for Cyclical Components of U.S. Time Series, 1970–2000^a

Variable (t)	Standard deviation		Cross-correlation with GDP (t - k)										
	absolute	relative/GDP	k: -4	-3	-2	-1	0	1	2	3	4		
GDP	0.34	1.0	-0.09	0.24	0.60	0.89	1.00						
Consumption	0.01	0.8	-0.24	0.03	0.34	0.64	0.84	0.87	0.74	0.54	0.27		
Investment	3.26	2.4	0.11	0.44	0.75	0.94	0.95	0.80	0.53	0.20	-0.10		
Cumulated inventories	2.35	1.8	0.74	0.89	0.88	0.69	0.35	-0.02	-0.32	-0.48	-0.48		
GDP deflator (level)	0.67	0.5	0.00	-0.16	-0.31	-0.42	-0.48	-0.49	-0.46	-0.42	-0.39		
CPI (level)	1.02	0.8	0.23	0.10	-0.07	-0.24	-0.41	-0.52	-0.56	-0.54	-0.49		
CPI (inflation)	1.29	1.0	0.48	0.59	0.63	0.56	0.38	0.15	-0.09	0.25	-0.31		
Stock prices	7.92	5.9	-0.50	-0.50	-0.37	-0.12	0.16	0.39	0.47	0.40	0.22		
Real estate prices	2.12	1.6	-0.18	-0.21	-0.16	-0.06	0.08	0.21	0.24	0.17	0.03		
Short-term rate nominal	1.31	1.0	0.38	0.56	0.68	0.67	0.50	0.21	-0.14	-0.44	-0.62		
Short-term rate real	1.11	0.8	-0.11	-0.03	0.07	0.14	0.15	0.07	-0.06	-0.22	-0.36		
Long-term rate nominal	0.82	0.6	-0.03	0.14	0.28	0.35	0.30	0.14	-0.07	-0.28	-0.41		
Yield spread	0.90	1.7	-0.51	-0.60	-0.63	-0.56	-0.39	-0.15	0.12	0.33	0.45		
Real effective exchange rate	2.96	2.2	0.08	0.11	0.08	0.00	-0.07	-0.12	-0.12	-0.08	-0.01		
DM-USD exchange rate	6.66	5.0	0.19	0.23	0.23	0.23	0.27	0.37	0.45	-0.42	0.27		
MI	1.78	1.3	-0.22	-0.23	-0.18	-0.08	0.05	0.16	0.22	0.24	0.22		
MB	0.87	0.7	0.25	0.37	0.42	0.39	0.28	0.12	-0.03	-0.13	-0.15		
Total loans	1.99	1.5	0.75	0.78	0.68	0.48	0.19	-0.11	-0.34	-0.45	-0.45		
			Cross-correlation with own (t + k)										
CPI (level)	1.02	0.8	0.38	0.61	0.81	0.95	1.00						
GDP deflator (level)	1.29	1.0	0.35	0.58	0.80	0.95	1.00						

Source: Agresti and Mojon (2003).

a. Standard deviation of and cross-correlation between the cyclical components of macroeconomic time series. The cyclical component was obtained from a band-pass filter BPF(6,40.8) a la Baxter and King (1999), and the data sources are fully described in Agresti and Mojon (2001, appendix 1).

1.2 Transmission Evidence from VARs

As noted earlier, we use the phrase monetary transmission mechanism to describe the effects of a change in the stance of monetary policy on real quantities and prices. In some cases we cite evidence from vector autoregressions (VARs) that have the interpretation of the response of different variables to an unanticipated shock to the implicit central bank reaction function. In other cases we refer to evidence embodied in traditional macroeconometric models maintained in the central banks. We recognize that, depending on one's preferred theory of monetary nonneutrality, one or another of the various pieces of evidence would be regarded as more relevant. We believe, however, that there is unfortunately not sufficient consensus over which model of nonneutrality is correct (or even most correct), and hence believe that a dogmatic approach of ruling out certain types of evidence would be unwise.

Our first set of evidence looks at VARs, drawing from previous research. We update these specifications to include current data (to see if that matters). For each area we consider three models. We first review the U.S. models and their results and then do the same for the euro area.

The first U.S. VAR follows the recursive identification procedure proposed by Christiano, Eichenbaum, and Evans (1999) that has become the benchmark in this literature. We analyze the variant proposed by Erceg and Levin (2002) that was designed to provide information on the composition of output responses to monetary shocks. Because of this focus, Erceg and Levin modified the Christiano, Eichenbaum, and Evans (1999) specification to include different components of GDP whose interest rate sensitivities might be expected to differ. Consequently, their model includes GDP and a host of demand components, along with a price deflator, a commodity price index, and the federal funds rate.

We depart from this by including only investment and consumption and using a slightly different commodity price series and the consumer price index (CPI) instead of the GDP deflator.⁴ We limit the demand components to consumption and investment because we do not have the further disaggregated data for the euro area anyway

4. There is no single commodity price series that is universally used in this literature. Our findings suggest that the choice of the series makes little difference to the estimated impulse responses, although whether the series is smoothed or not makes a slight difference in reducing the size of the "price puzzle" discussed below.

(and we favor treating both areas symmetrically). But even with this crude separation, we can study the composition of the adjustment that underlies the output responses. Given this aim we also replace GDP with GDP less the sum of consumption and investment (that is, by net exports and government spending, which we call the rest of GDP henceforth). This substitution provides us with a parsimonious way to show both the total GDP response to monetary shocks (obtained as the sum of the responses of consumption, investment, and the rest) and its composition. Moreover, this procedure can be interpreted as a quick way to impose in the VAR the constraint provided by the national accounting identity, of the type usually imposed in traditional macroeconomic models. As our choice does not lead to overall GDP responses to monetary shocks that differ from previous findings, we are confident that we are not badly misspecifying the model by making this choice. We make this same substitution in all of the other VARs.

For our consumption series we use private consumption, that is, the sum of nondurable goods, services, and durable goods consumption. For investment we use total private sector investment. These aggregates are the closest match for GDP components that are available for the euro area: private consumption and private investment.⁵

Our baseline estimation period for the U.S. sample begins the first quarter of 1960 and ends in the fourth quarter of 2001—the starting date is given by the availability of the official data for the money supply figures and the ending date by the last quarter with data that were not preliminary as of the time we began the analysis. However, we also consider another subsample that runs from 1965 to the fourth quarter of 2001, but omits the data from the fourth quarter of 1979 until the fourth quarter of 1983. The 1965 start-date is chosen because this is when the market for federal funds began to operate in its current format. The excluded period covers the interval when the Federal Reserve's operating procedures changed to emphasize the importance of nonborrowed reserves.⁶ Finally, we also look at a sample that runs from 1984 to the fourth quarter of 2001. This covers the

5. In the case of the euro area, we are missing an exact deflator for euro area government investment because the ESA 95 system of national account does not require total investment to be broken down into its private sector and public sector components. See the data appendix for an explanation of the construction of private investment series for the euro area, Germany, France, Italy, and Spain. However, for the VARs where it is possible to experiment with both private and total investment, there are no important differences that depend on which of these series is used.

6. See Bernanke and Mihov (1998) and Christiano, Eichenbaum, and Evans (1999) for a discussion of the changes in the Federal Reserve operating procedures.

most recent part of the sample only and spans the period during which the operating procedures were relatively stable. The models are estimated with four lags for the first two samples and, in order to preserve degrees of freedom, with two lags for the 1984–2001 sample.

Our second model is based on an identification procedure proposed by Gordon and Leeper (1994). Their model adds a long-term (ten-year) interest rate and M2 to the list of variables examined by Erceg and Levin. Gordon and Leeper opt for an alternative set of identifying restrictions that focus on the information set that the central bank could be expected to have at the time it was setting the short-term interest rate. Accordingly, they do not allow contemporaneous data on inflation and GDP to influence this decision—leaving only contemporaneous commodity prices, the long term interest rate, and M2 as potentially affecting the contemporaneous federal funds rate. In contrast, contemporaneous prices and GDP components enter the money demand equation. Our decomposition of the demand components leads naturally to modifying this identification strategy by assuming that the innovations of consumption, investment, and the rest of GDP have no effect on the innovation of the federal funds rate while they have an effect on the innovation of M2.

Our third model is taken from Christiano, Eichenbaum, and Evans (2001). This model includes consumption, investment, GDP, the CPI, a real wage variable, a labor productivity measure, real corporate profits, the federal funds rate, M2 growth, and the Standard and Poor's 500 stock price index deflated by the CPI. We substitute private consumption and investment for the consumption and investment series that they use in order to match the euro area data (where disaggregated figures are not available).⁷ Given the substantial difference between this specification and the other two VARs we consider, this alternative particularly important.

Turning to the results, most of our main findings (aside from the composition of the output response) are summarized in figure 1, with each of the three panels describing one of the models. The Christiano, Eichenbaum, and Evans and the Erceg and Levin models are each just identified, so that the procedure for computing confidence intervals for impulse responses is easily implemented (Sims and Zha, 1999). The

7. In Christiano, Eichenbaum, and Evans (2001), consumption is defined as the sum of nondurable goods, services, and government consumption, while the investment they include in their VAR is the sum of gross private sector investment and durable consumption. We thank Larry Christiano and Charlie Evans for providing us their data.

graphs report the point estimate of the impulse response and the confidence band formed by tenth and the ninetieth percentiles based on 1,000 Monte Carlo simulations for twenty quarters after the initial shock.⁸ In the case of the Gordon and Leeper model, which is over-identified, the point estimates and error bands, again the tenth and the ninetieth percentiles of the simulated impulse responses, are based on the Bayesian procedure advocated by Sims and Zha (1999).⁹ We notice that the responses of consumption and investment estimated with this procedure are more persistent than—and about twice as large as—the one obtained with the other two VARs.

As a matter of course, the confidence intervals for the second half of the sample are much wider, so these results are in general less certain. But, despite the substantial differences across the VAR specifications, two consistent findings emerge from our analysis of monetary policy shocks. First, the impulse responses clearly show that following an innovation in the funds rate, output declines within one or two quarters and reaches its peak decline within four to eight quarters.¹⁰ The responses are such that the decline is significantly different than zero around the peak (and this is true even for the short sample). The standard errors grow as the horizon extends beyond two years, so that precise statements are not warranted, but we cannot reject the proposition that output is back at its baseline five years after the shock in almost all of the cases.

The second consistent finding is that price responses are more sluggish than the output responses. Here the exact shapes are somewhat sample and model specific. In all of the specifications and time periods, prices show little change in the first couple of quarters after the monetary policy disturbance. In some of the specifications, prices actually rise for more than a year after an increase in interest rates. Sims (1992) labeled this perverse price response the “price puzzle” and explained it as possibly reflecting omitted variables from the VAR to which the Federal Reserve might be responding. Subsequently, Christiano and Eichenbaum (1995), Barth and Ramey (2002), and others have suggested the possibility that this could be due to the effect of higher interest rates on firms’ short-run financing costs. For

8. All the simulations were performed with Rats 5.0. The original Rats program for computing error bands was modified to report percentiles of the simulated impulse responses instead of adding multiples of the standard errors to the mean of the simulated impulse responses.

9. We thank Jennifer Roush for assistance in implementing the Bayesian procedure and computing these confidence intervals.

10. The output responses are always recovered by summing the components.

our purposes, explaining this phenomenon is less relevant than noting that the slow response of prices to policy shocks seems to be a pervasive feature of the data.

In the long baseline sample, the estimated responses after the first year are more in line with standard theoretical predictions. In both the Erceg and Levin and the Christiano, Eichenbaum, and Evans models, the price declines are estimated to be significant by eight to twelve quarters after the initial shock. After that, while the uncertainty surrounding the point estimate becomes fairly large, we typically cannot reject the hypothesis that the price level eventually settles down to a new permanently lower level (with no long-run effect on inflation). For the Gordon and Leeper model, the price-level response even in the long sample is almost always indistinguishable from zero.

In the other two samples, and particularly the recent subsample, we cannot in general detect any statistically significant price effects from the change in monetary policy. In most of these cases, even the point estimates suggest weak responses. Thus, we conclude that the VAR evidence on the transmission mechanism for the United States is much less clear regarding prices than output.

Turning to the euro area, we start with an area-wide analysis, using synthetic data that is created by combining country-level macroeconomic variables to form aggregate data for the area as a whole. The first model we consider for the area-wide analysis follows the specification proposed by Peersman and Smets (2003) and includes GDP components, the harmonized index of consumer prices (HICP), M3,¹¹ the money market interest rate, and the effective exchange rate of the euro as endogenous variables. In addition, the model includes three U.S. variables that account for shocks to the world economy: the index of commodity prices already used in the VAR models of the United States described above, U.S. GDP, and the federal funds rate. These three variables are exogenous. The monetary policy shock is identified by a Cholesky decomposition, with the variables ordered as above. We report estimates for two samples: 1970–2000, the longest available sample period, and 1980–2000, which starts with the beginning of the European Monetary System (EMS).¹²

11. M3 is the natural choice among monetary aggregates given the importance it has in the monetary policy strategy of the European Central Bank.

12. Within the EMS, countries that then belong to the European community, namely, Belgium, France, German, Italy, Luxembourg, and the Netherlands, pegged their currency to the ECU, a basket of their currencies. De facto, currencies were pegged to the Deutsche-Mark in order to import the credibility of the Deutsche Bundesbank.

Figure 1a. U.S., Ercerg-Levin Identification (quarterly data)

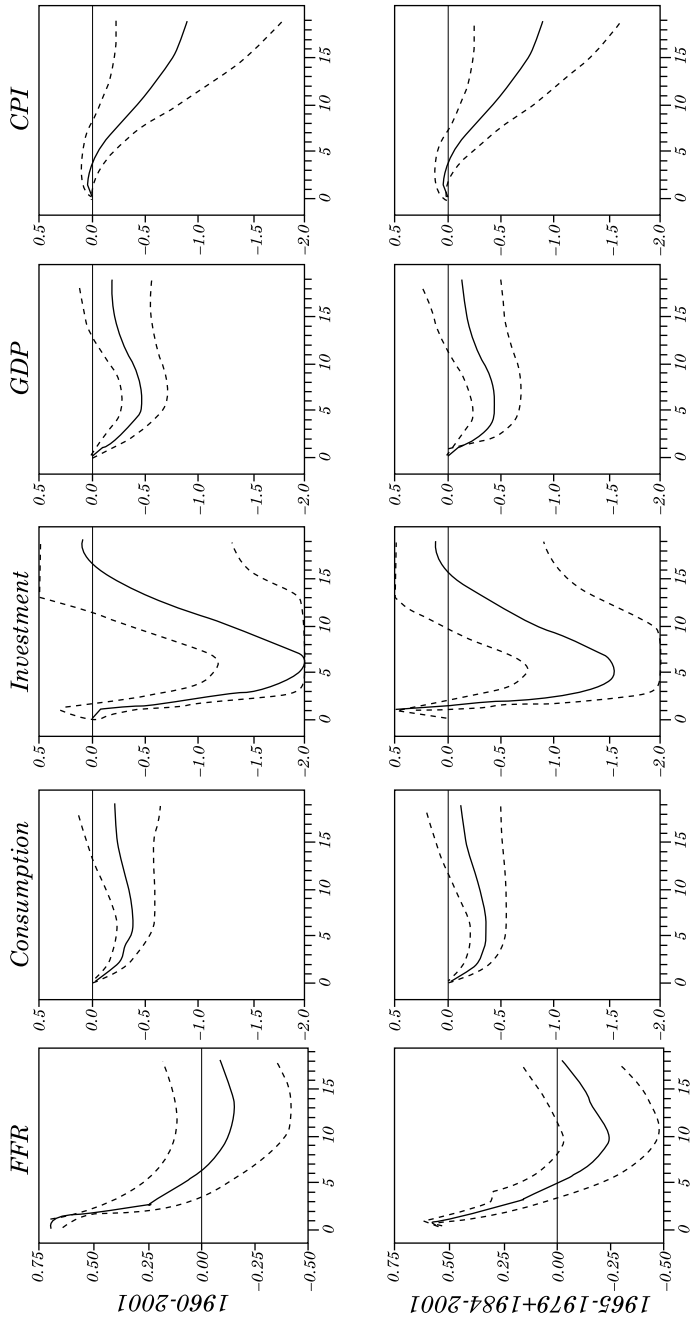


Figure 1a. (continued)

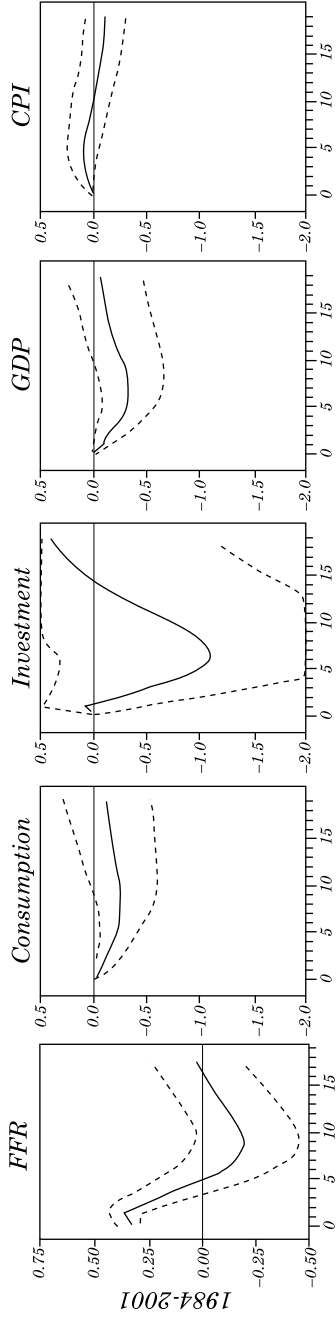


Figure 1b. U.S., Gordon-Leeper Identification (quarterly data)

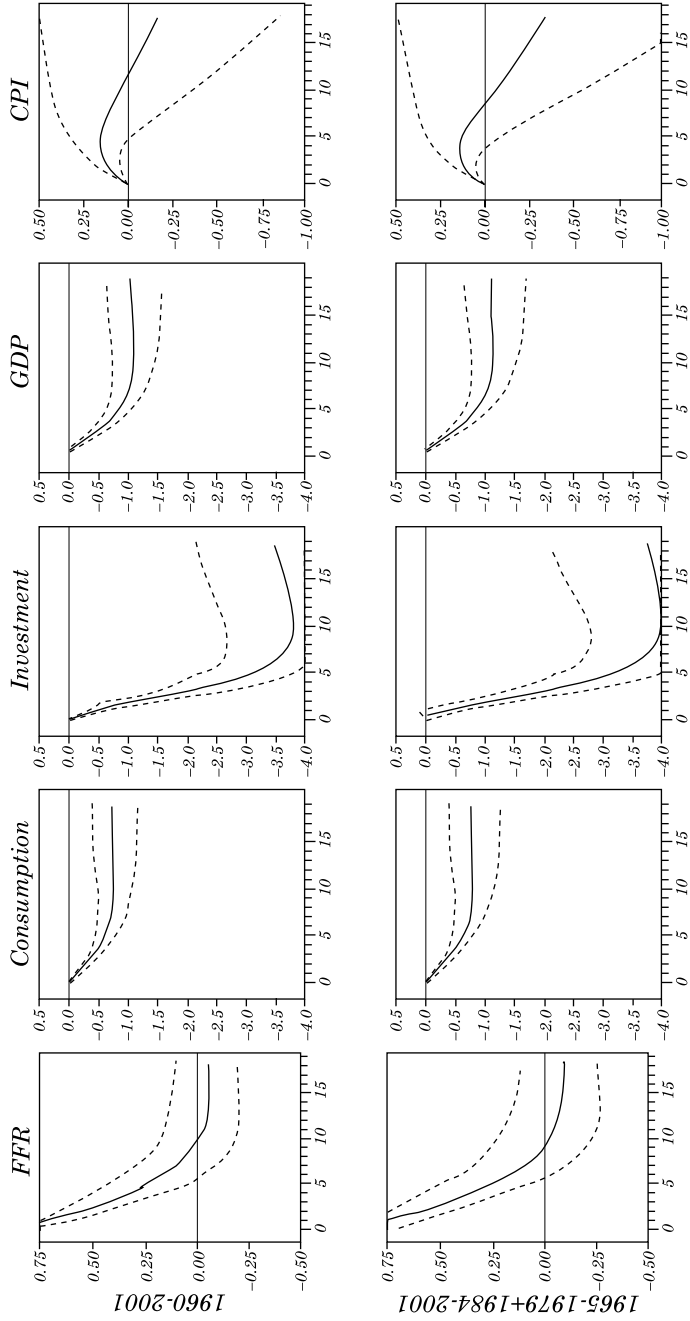


Figure 1b. (continued)

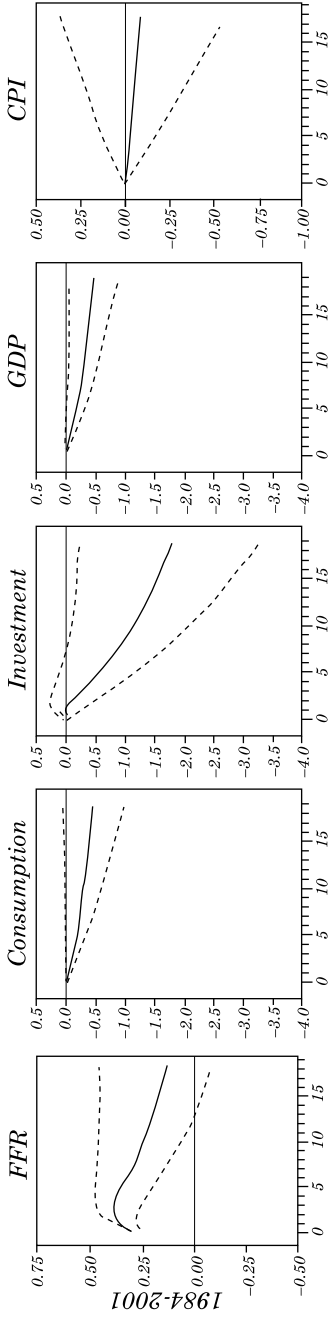


Figure 1c. U.S., Christiano, Eichenbaum, and Evans Identification (quarterly data)

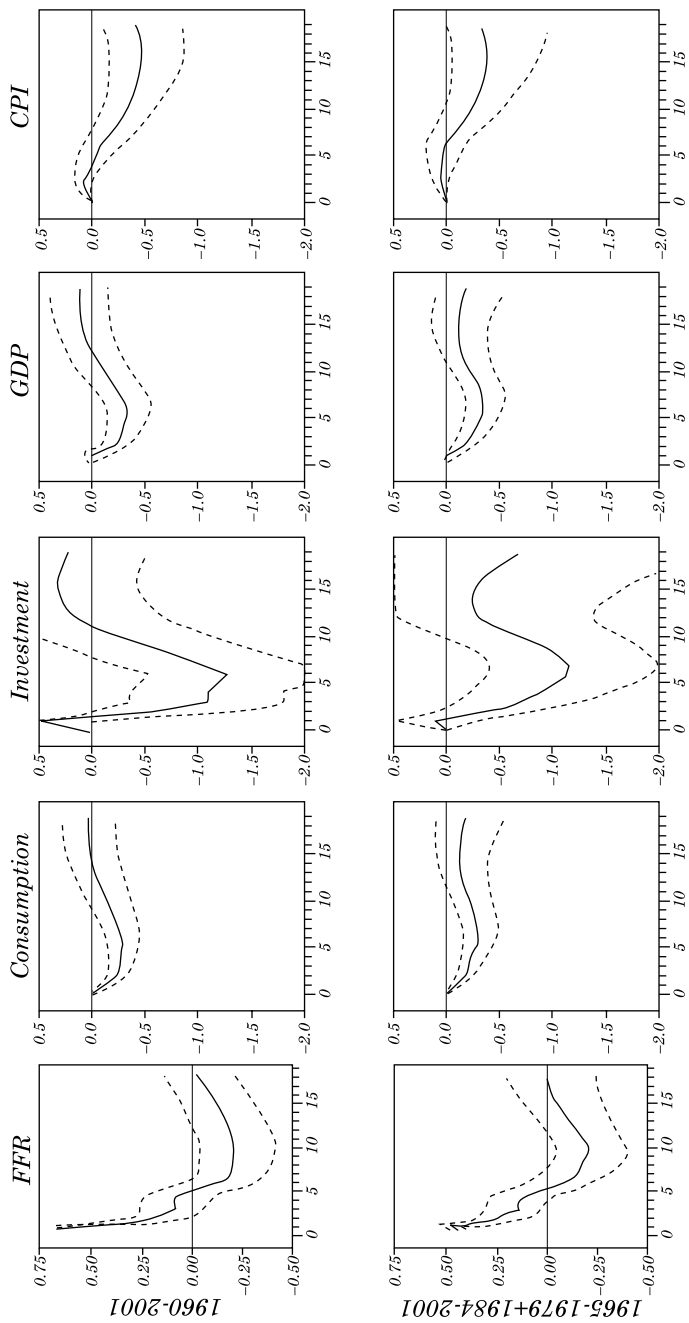
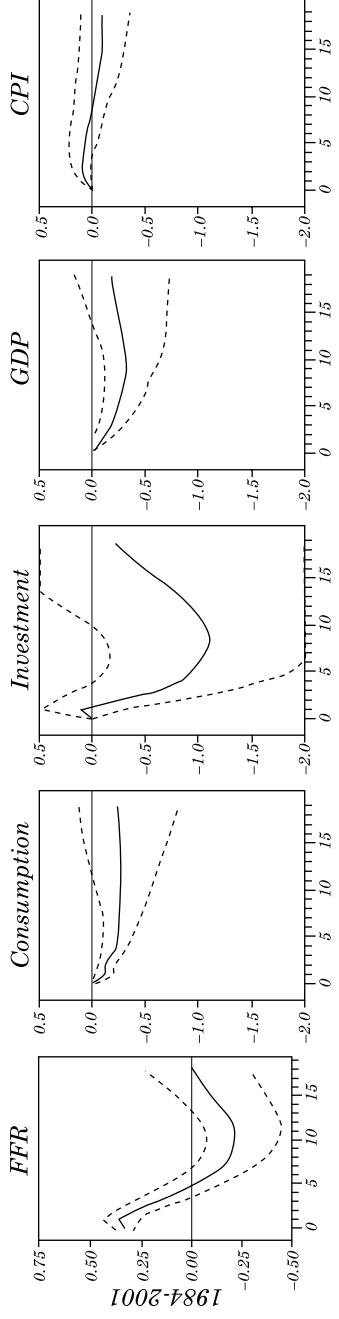


Figure 1c. (continued)



We also report a second version of the Peersman and Smets (2003) model without M3. We consider this alternative for two reasons. First, monetary aggregates were not as prominent in the European central banks' monetary policy strategy in the 1970s as they subsequently became. Second, euro area synthetic monetary aggregates have only recently been backdated to the 1970s. Our models that include M3 for the 1970s should then be taken with caution, at least until the econometric properties of this new series are better known.

Our third model mimics Christiano, Eichenbaum, and Evans (2001) for the euro area.¹³ To avoid a perverse money response for one of the two samples, we need to substitute the stock price index by the real effective exchange rate within the model. However, this substitution does not change the effects of monetary shocks on other variables of the model. All the specifications that we analyze also include the time trend and other exogenous variables that Peersman and Smets (2003) advocate.

In order to maximize the degrees of freedom, all the results presented here are based on models estimated with two lags.¹⁴ In addition, the consumer price indices and the monetary aggregates are entered as growth rates. This transformation improves the stability of some of the impulse responses.

Figure 2 summarizes the main findings of the three VARs which we estimated using euro area synthetic data. The output and price responses to the identified monetary policy shock are quite similar to what is observed for the United States. In particular, the response of output to the monetary policy shifts is hump shaped, with the peak occurring about one year after the shock. Likewise, the response of prices is more gradual than that of output. Finally, the effects on output and on inflation are temporary.

However, in contrast with the U.S. estimates, the uncertainty of the responses does not fall when the sample is extended prior to 1980. This is one indication of the instability amongst these European economies in the 1970s.

As a robustness check we also analyze a similar set of VARs for France, Germany, Italy, and Spain, which together account for 80 percent of the euro area GDP. Our goal in doing so is to verify that the use of the synthetic data is not masking any obvious patterns

13. The additional variables relative to the Peersman and Smets model are productivity, profits, and workers' compensation.

14. The pattern of responses are, however, quite similar with either three or four lags in most cases.

Figure 2a. Euro Area, Peersman-Smets Identification with M3 (quarterly data)

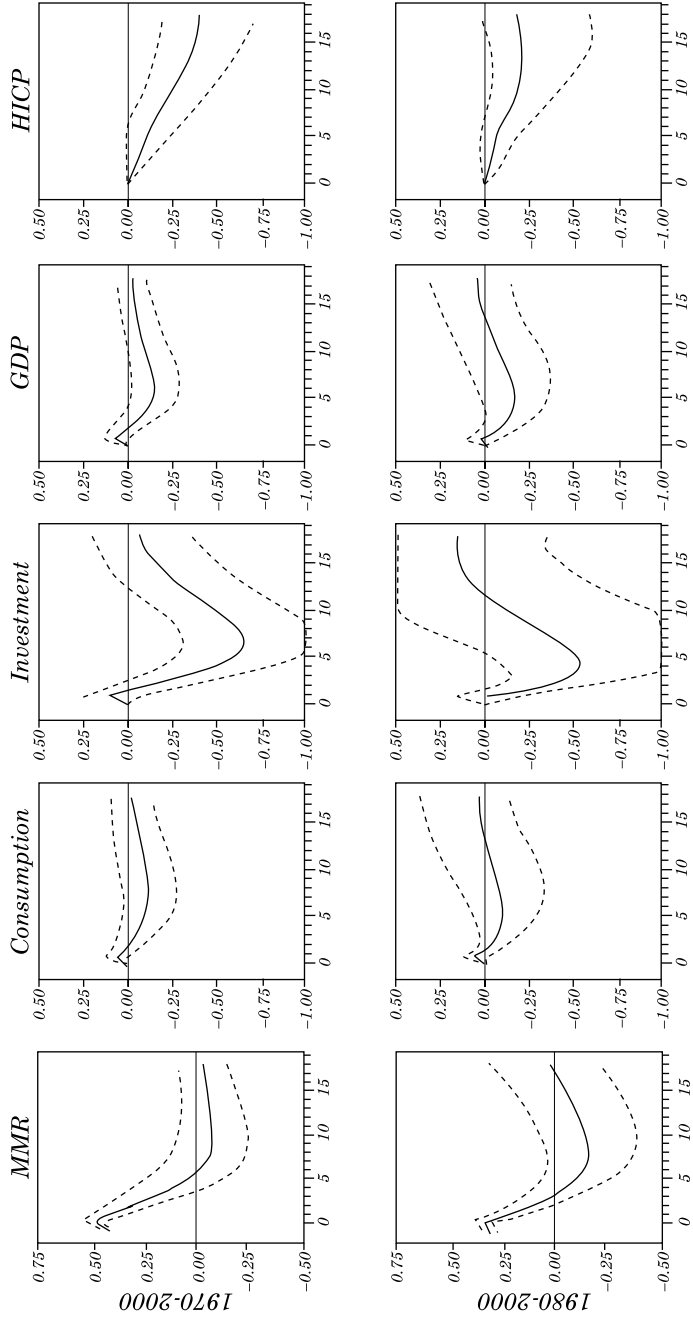


Figure 2b. Euro Area, Peersman-Smets Identification without M3 (quarterly data)

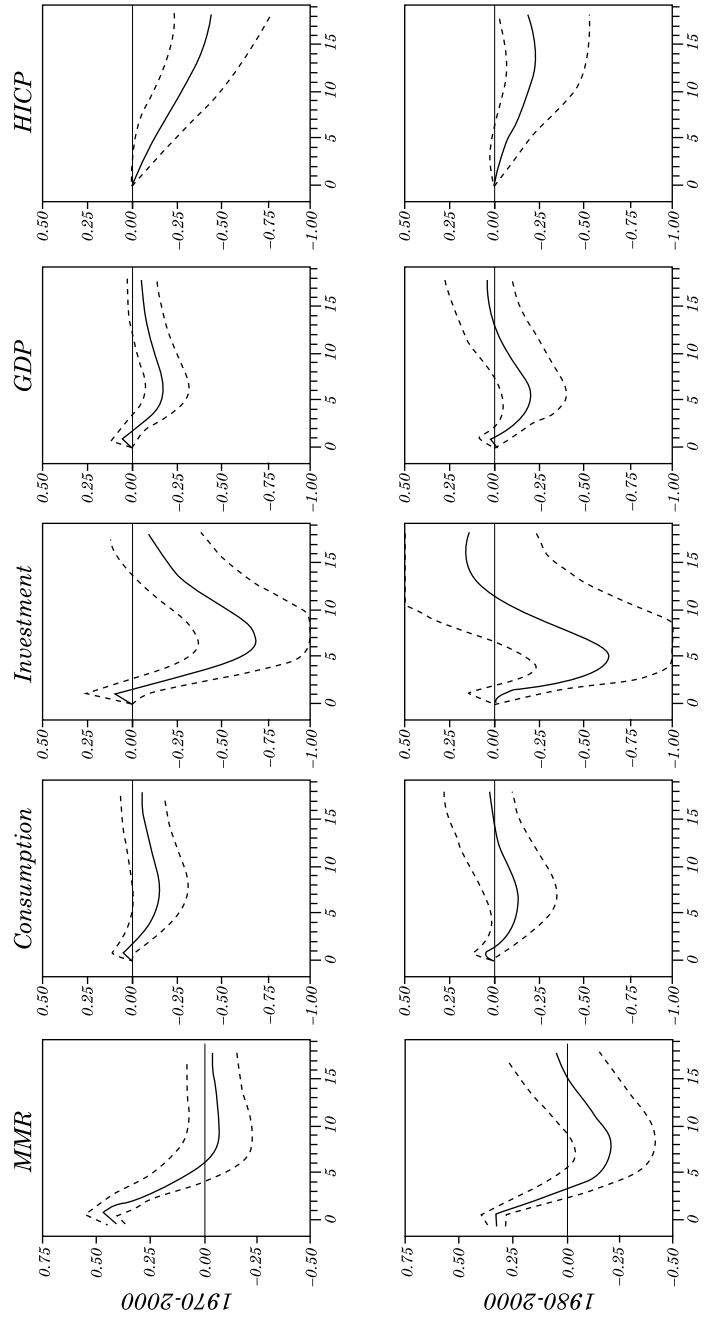
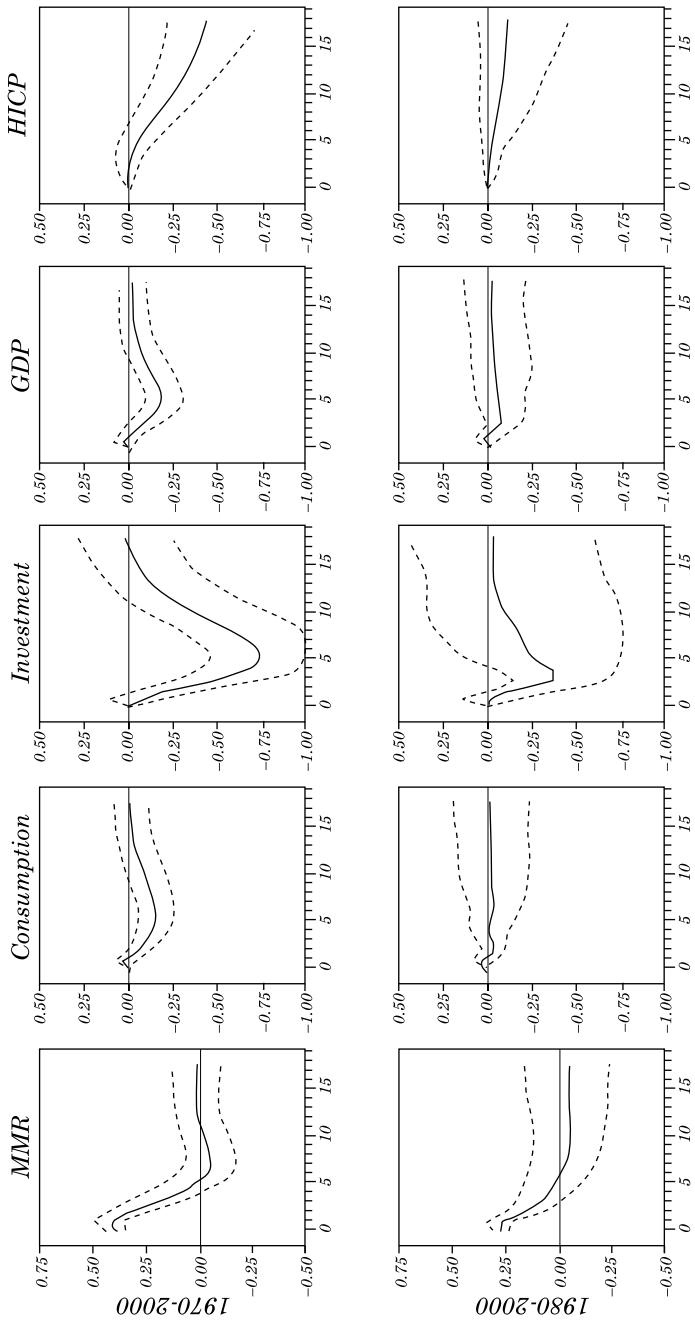


Figure 2c. Euro Area, Christiano, Eichenbaum, and Evans Identification (quarterly data)



that would be present at the country level. To do so, we update the Mojon and Peersman (2003) VARs for these four countries. We include in the VAR the breakdown of GDP into its main components as was done for the United States and with the euro area synthetic data.¹⁵

The sample period runs from the first quarter of the 1980, right after the European Monetary System started, to 2001—so that it coincides with our short sample for the model estimated with euro area synthetic data. However, given all of the shocks that hit the EMS, we recognize that identifying monetary policy shocks in this short sample is difficult. A full set of robustness checks for these results would take us too far astray. But because the findings are in line with the more comprehensive analysis conducted with Mojon and Peersman (2003), we believe that they are representative of what a typical VAR based approach suggests about the transmission mechanism in these countries. Thus, we see these results as another independent way to check whether our findings with euro area synthetic data are accidental.

The results are shown in figure 3, with one panel for each country. In general the country-level results are qualitatively similar to the findings for the area as a whole, but quantitatively the responses of consumption and GDP are even weaker than in the area-wide data and are almost never significantly different from zero. In the case of Germany, consumption remains above baseline for three quarters after the initial shock. Also, investment appears less persistent at the country level than at the euro area level—although these responses typically are significant after the first year. Finally, prices adjust gradually downward in Italy, Spain, and France, but they hardly deviate from the baseline in Germany.

Overall, we read the evidence from the countries as confirming the area-wide findings and showing that both are broadly consistent with the consensus view on the effects of monetary policy in the United States.

1.3 Transmission Estimates from Large-scale Models

We now look at an alternative characterization of the monetary transmission, that provided by large-scale “structural” macroeconomic models. Relative to VARs, these models incorporate vastly different

15. Two other differences with Mojon and Peersman are that we use private investment instead of total investment and that we extend the sample period to include the first three years of the monetary union. See the appendix for further details.

information sets and modeling priors, hence a rigorous comparison may look impossible. Nonetheless, it is precisely this difference that we regard as potentially informative. If each of these two sets of models incorporate, to some extent, essential features of the data and of the correctly identified transmission mechanism, then findings that are robust across the two may be particularly reliable, as they do not depend on arbitrary modeling choices. In this sense, after having examined several benchmark VARs, we view the contrast between these and structural models as more informative, at the margin, than further comparisons among alternative VARs.

We consider two sets of model results. The first, for the United States, comes from simulations of the Federal Reserve Board's macroeconomic model of the U.S. economy (FRB/US).¹⁶ Ludvigson, Steindel, and Lettau (2002) report some comparisons of how policy rate changes in this model compare to predictions made by the Washington University macroeconomic model and the Data Resources, Incorporated model. Along the dimensions that we emphasize, it appears that these three models are relatively similar.

The euro area results are obtained from two sources. The first is a euro-area-wide model (AWM) developed by the ECB staff (Fagan, Henry, and Mestre, 2001; Dieppe and Henry, 2002), estimated on synthetic data. The second is an aggregation of results from national models developed by the national central banks (NCBs; see van Els and others, 2001). These findings are built up from a set of simulations of identical monetary shocks in each country (in which the intra-area exchange rates are fixed). Likewise a harmonized treatment of long-term interest rates and exchange rates was imposed. Thus, the simulation is intended to crudely approximate the conditions that would prevail in a currency union.

The specific interest rate path that is considered is an eight-quarter increase in the money market rate (the fed funds rate in the U.S. case) by 100 basis points. The long-term interest rate and the exchange rates were respectively assumed to move according to the expectations hypothesis and the uncovered interest parity condition. Specifically, the exchange rate initially appreciates by 2 percent and then gradually returns to baseline over two years; the long-term rate adjusts up immediately, by about 20 basis points, and gradually returns

16. We thank Flint Brayton and Chris Erceg for providing these results to us. The simulations are run with the standard version of the model in which expectations are based on VAR forecasts; see Reifschneider, Tetlow, and Williams (1999) for a full description of the model and its properties.

Figure 3. Euro Area, Country Level VARs (quarterly data)

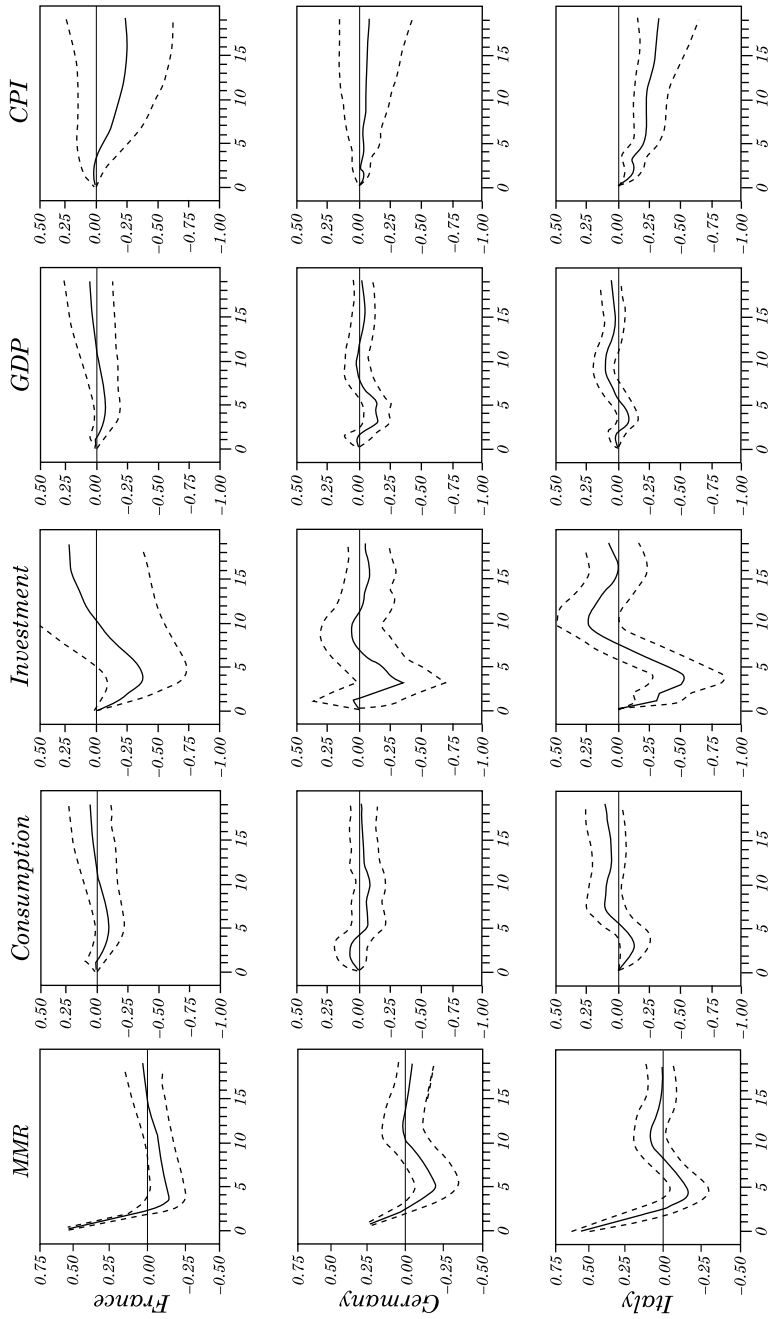
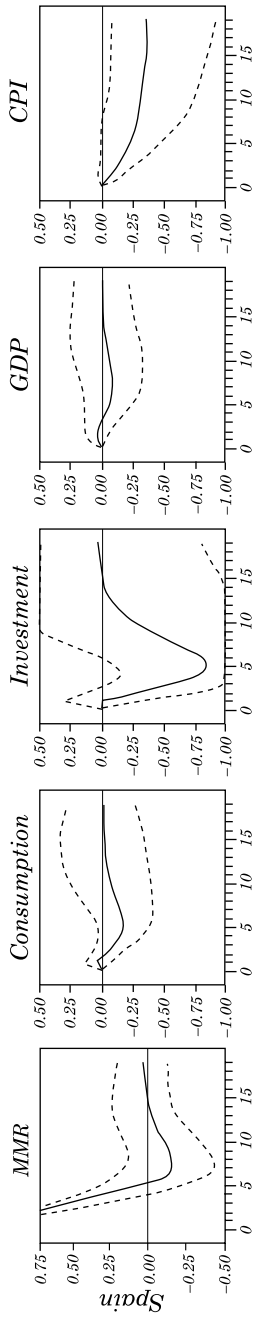


Figure 3. (continued)



to baseline. While the nature of the experiment conflicts with the Lucas policy regime invariance criterion (since the model coefficients are assumed unchanged), we still believe that it is informative for the small, temporary shock that is envisaged.

The left panel of table 3 reports results on the United States.¹⁷ These results are quite similar to those obtained from the VARs in terms of the reactions of prices, output, and the components of output. In particular, output and consumption responses are hump shaped with a maximum decline at the beginning of year three, while investment keeps falling all the way through the third year. Prices are virtually unchanged for the first four quarters after the tightening. From year one onward, prices fall steadily for the next two years. Thus, the relatively slower response of prices compared to output that was observed in the VARs is also present in the FRB/US simulations.

The right-hand side of table 3 reports the euro area simulations. Again, despite the methodological differences, the effects on output and on prices are qualitatively similar to the outcome of the VAR models of the euro area. The hump-shaped response of GDP (which begins moving back to the baseline from year four in the area-wide model) and the gradual response of prices also matches the results obtained for the United States. Robustness across models may suggest that the results reflect underlying features of the data. Moreover, these results are broadly consistent with the pattern observed at the national level in the simulations based on the NCBs model, at least in qualitative terms.¹⁸

2. EVIDENCE ON THE COMPOSITION OF OUTPUT RESPONSE

The composition of the output response has attracted much less attention than the size and timing of the overall GDP and price responses discussed above (with the notable exceptions of Bernanke and Gertler, 1995, and Erceg and Levin, 2002). Yet, whether consumption or investment responds more, or more quickly, to a monetary tightening is an issue of clear importance in the policy debate and in welfare analyses.

17. The results we describe here are very close to the ones (not reported) obtained when following an initial shock, the funds rate evolves according to a Taylor rule, that is, so that it depends on the gap between inflation and the target rate of inflation and the output gap.

18. For a detailed presentation of these results, see van Els and others (2003).

Table 3. Effects of Monetary Policy Shocks in Large-scale Models
Deviation from baseline

Variable	Region and model											
	United States			Euro area						AWM		
	FRB-US			NCBs			Euro area			AWM		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Short-term interest rate	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00
Long-term (ten-year) interest rate	0.16	0.06	0.00	0.16	0.06	0.00	0.16	0.06	0.00	0.16	0.06	0.00
Effective exchange rate	1.60	0.63	0.00	1.60	0.63	0.00	1.60	0.63	0.00	1.60	0.63	0.00
CPI	-0.07	-0.41	-1.01	-0.09	-0.21	-0.31	-0.13	-0.26	-0.37	-0.13	-0.26	-0.37
GDP	-0.35	-1.28	-1.37	-0.22	-0.38	-0.31	-0.27	-0.71	-0.79	-0.27	-0.71	-0.79
Consumption	-0.37	-1.35	-1.44	-0.12	-0.23	-0.19	-0.19	-0.57	-0.63	-0.19	-0.57	-0.63
Investment ^a	-0.31	-1.79	-3.16	-0.34	-1.04	-1.22	-0.72	-2.29	-2.93	-0.72	-2.29	-2.93
Disposable income	-0.03	-0.49	-0.67	-0.03	-0.16	-0.14	-0.16	-0.55	-0.65	-0.16	-0.55	-0.65
Unemployment	0.12	0.66	0.77	0.04	0.11	0.17	0.08	0.36	0.61	0.08	0.36	0.61

Source: For the United States; private correspondence with Flint Brayton; Euro area NCBs: van Els and others (2003); Euro area AWM: Dieppe and Henry (2002).
a. Private sector investment in the case of the United States and total investment for the euro area.

To measure the composition of the output response, we take the ratio between the (monetary policy induced) change in each demand component and the total change, obtained as the sum of the changes of the various components.¹⁹ In particular, we focus on consumption and investment, computing what we term their *contributions* to the response of the private sector domestic demand (PSDD)—the sum of consumption plus investment. We view this normalization as a way to minimize the importance of the shortcut that we took in modeling the rest of GDP in the VARs. Also, it allows a direct comparison with the results obtained in the dynamic stochastic general equilibrium models assessed in the next section of the paper, where only consumption and investment are modeled.²⁰ In what follows we consider cumulative changes, in order to smooth out some of the noise that can be present in the responses (particularly in the first periods).²¹ Despite this smoothing, the estimated contributions in these first few periods are rather unstable in a few cases. This occurs whenever the overall response of PSDD to the monetary policy shock is initially close to zero.

A major advantage of the contribution measures is that they are unit-free statistics that can be compared across models and countries, thus sidestepping the problems of comparability among VARs and structural models. This is because, by focusing on a comparison of how much investment or consumption move relative to PSDD following a given policy shift, the nature of the shift that moves both the components to be compared is, in general, less relevant. One exception to this is when the persistence of the policy shift is significantly altered. However, this is unlikely to be the case for the kind of shifts that are considered throughout the paper.

19. If the model is specified in a log-linear form, we recover the contribution as follows: we first take the ratio of the responses of the consumption and investment to the response of GDP, each relative to baseline (these are then semi-elasticities); we multiply these two ratios by the shares of consumption and investment in GDP, respectively; we normalize the results so that they add up to one. In particular, for the euro area we used the average consumption and investment shares over the 1970 to 2000 period: 0.60 and 0.186, respectively. For the United States, we used the average shares from 1960 to 2001: 0.66 and 0.15, respectively.

20. Given that in this metric, the contribution of investment and the contribution of consumption add up to 1, we report only the contribution of consumption for the sake of space.

21. Note that cumulating up to time t the responses to a one-off shock occurring in $t - k$ can also be interpreted as observing, at time t , the response to a shock sustained from $t - k$ to t ; the latter is the measure we adopt when looking at structural macroeconomic models.

In the upper half of table 4, we report the estimated contributions based on the U.S. VAR models. The table shows the median contribution along with the tenth and ninetieth percentiles of 1,000 Monte Carlo simulations. In the lower half of the table, we report the point estimates for the FRB/US model. Table 5 reports analogous figures for the euro area VARs and structural models.

Rather than discuss the many potential comparisons between the table 4 and table 5 estimates, we combine the simulations from the different VARs to form one complete set of estimated contributions for each economy. This means that the U.S. distribution is based on 9,000 simulated draws (three models, over the three samples), while the euro area distribution is based on 6,000 simulated draws (three models, over two samples). The three panels in figure 4 show the pair of distributions at three horizons (quarters four, eight, and twelve). On each of the distributions, we also draw vertical lines to show the point estimates from the large-scale models.

Figure 4 provides the basis for our assertion that there is an output composition puzzle. It is apparent from the figure that the size of the consumption contributions in the two economies is quite different. The difference is significant in both economic and statistical terms. For instance, focusing on the VARs, one would conclude that the difference in the medians of the distributions is 32 percentage points at four quarters and remains above 13 at twelve quarters. A formal Kolmogorov-Smirnov test for the equality of two distributions rejects the hypothesis of equality (at a significance level well below one percent at each of the three horizons).

Another way to see the large difference between the VAR estimates for the contributions is to examine the cumulative distributions of these data. At the four-quarter horizon, more than two-thirds of the euro area simulated consumption contributions are below 0.4. In contrast, only about 5 percent of the U.S. simulated contributions are below 0.4. At the twelve quarter horizon, 86 percent of the simulated euro area consumption contributions are below one-half, while only 41 percent of the U.S. contributions are below one-half.

Importantly, these large differences are not tied to using VARs; they are also apparent in the implied contributions coming from the large-scale models. The FRB/US model implies much larger consumption contributions than do the U.S. VARs and all euro area structural models. For instance, the point estimates from the FRB/US model and, for the euro area, the aggregation of the national models consistently show differences in consumption contributions on the order of 30 percentage points.

Table 4. United States: Contribution of Consumption to the Private Sector Domestic Demand Response to a Monetary Policy Shock
Percentiles and point estimates

VAR model	Horizon	Sample of estimation and percentile ^a											
		1960:1 to 2001:4				1965:1 to 1979:3 + 1984:1 to 2001:4				1984:1 to 2001:4			
		10th	50th	90th	90th	10th	50th	90th	90th	10th	50th	90th	
Erceg and Levin (2002)	4 quarters	0.55	0.69	0.96	0.85	0.52	0.64	0.85	0.85	0.11	0.58	1.45	
	8 quarters	0.47	0.55	0.65	0.62	0.45	0.53	0.62	0.33	0.50	0.90		
	12 quarters	0.48	0.56	0.66	0.63	0.46	0.54	0.63	0.34	0.52	0.86		
Christiano, Eichenbaum, and Evans (2001)	4 quarters	0.54	0.68	0.97	0.80	0.49	0.61	0.80	0.47	0.69	1.30		
	8 quarters	0.47	0.56	0.69	0.65	0.46	0.54	0.65	0.46	0.56	0.79		
	12 quarters	0.48	0.58	0.72	0.65	0.46	0.55	0.65	0.45	0.53	0.71		
Gordon and Leeper (1994)	4 quarters	0.42	0.51	0.62	0.66	0.44	0.54	0.66	-1.34	-66.7	0.77		
	8 quarters	0.40	0.48	0.55	0.56	0.41	0.48	0.56	-0.11	0.53	1.24		
	12 quarters	0.39	0.47	0.54	0.55	0.39	0.47	0.55	-0.02	0.55	1.03		
Federal Reserve Board—U.S. model (point estimate)	4 quarters		0.81										
	8 quarters		0.74										
	12 quarters		0.66										

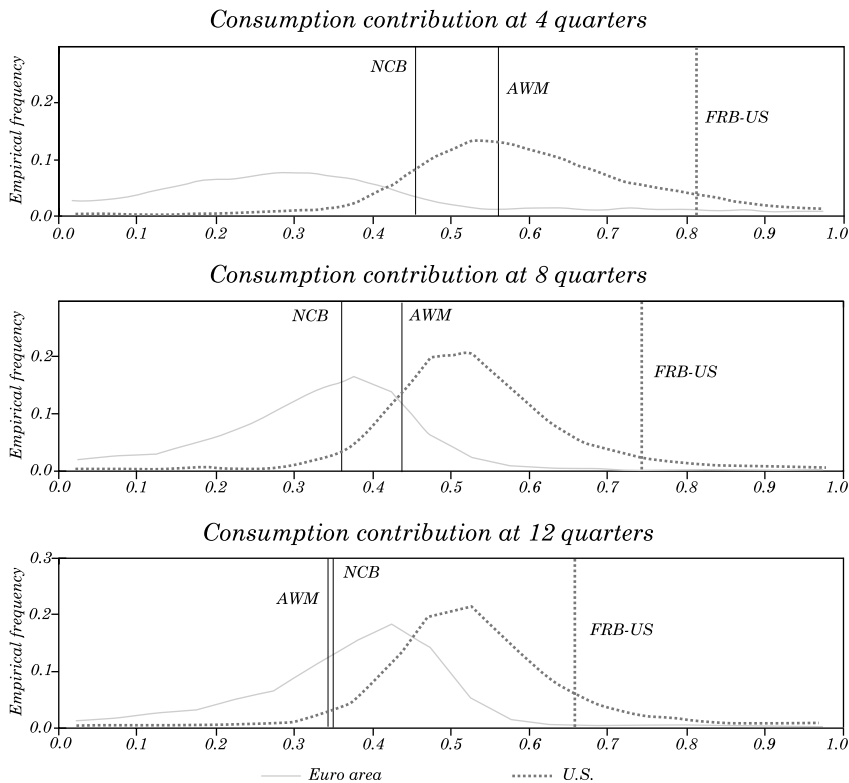
a. Percentiles are based on 1,000 Monte Carlo simulations.

Table 5. Euro Area: Contribution of Consumption to the Private Sector Domestic Demand Response to a Monetary Policy Shock
Percentiles and point estimates

VAR model	Horizon	Sample of estimation and percentile ^a								
		1980:1 to 2000:4			1970:1 to 2000:4					
		10th	50th	90th	10th	50th	90th	10th	50th	90th
Peersman and Smets (2003) baseline model	4 quarters	-0.79	0.20	0.84	-0.71	0.50	1.77			
	8 quarters	0.04	0.36	0.48	-0.24	0.31	0.52			
	12 quarters	0.15	0.43	0.58	-0.07	0.35	0.51			
Peersman and Smets (2003) without M3	4 quarters	-0.70	0.19	1.01	-0.67	0.49	1.82			
	8 quarters	0.04	0.35	0.46	-0.10	0.35	0.53			
	12 quarters	0.15	0.43	0.53	0.04	0.38	0.52			
Christiano, Eichenbaum, and Evans (2001)	4 quarters	-0.43	0.24	0.57	-0.19	0.28	0.45			
	8 quarters	-0.34	0.27	0.69	0.21	0.35	0.44			
	12 quarters	-0.17	0.36	0.89	0.23	0.37	0.46			
Central bank large-scale models (point estimates) National models (NCBs)	4 quarters		0.45							
	8 quarters		0.36							
	12 quarters		0.35							
Area-wide model (AWM)	4 quarters		0.57							
	8 quarters		0.43							
	12 quarters		0.34							

a. Percentiles are based on 1,000 Monte Carlo simulations.

Figure 4. Distributions of the Consumption Contribution in the VARs and the Structural Models in the Euro Area and the U.S.



Note: The density curves are based on 6,000 simulations for the euro area (1,000 draws for each of the two samples for each of the 3 VARs) and 9,000 simulations for the U.S. (1,000 draws for each of the three samples for each of the 3 VARs). The vertical lines indicate the contributions as obtained by the structural models.

Given its structural nature, for the FRB/US model it is relatively easy to understand why consumption adjustments are so important. A key part of the transmission mechanism in the model is that changes in the federal funds rate move long-term rates that lead to changes in the value of the stock market. Consumption is estimated to strongly respond to the change in wealth (see Reifschneider, Tetlow, and Williams, 1999). These wealth effects are also quantitatively significant in the Washington University macroeconomic model and the Data Resources, Incorporated model. To the contrary, the effect of stock

market prices on wealth and subsequently on consumption is not a prominent feature of the structural models for the euro area (see van Els and others, 2003).

As a further cross-check against figure 4, we also compute the consumption contributions implied by the VARs for France, Germany, Italy, and Spain. The top part of table 6 displays the contributions (median, tenth percentile, and ninetieth percentile) that correspond to the VAR results shown in figure 3. The lower part shows the contributions from the country-level structural models—together with similar calculations for the smaller countries in the euro area, these aggregate to the “NCBs” findings shown in tables 3 and 5.

The noise in the underlying VARs carries over to the contribution statistics, so the individual confidence intervals in table 6 are wide. But when we combine the results from the four countries a clearer picture emerges. Figure 5 shows this combined distribution along with the one from figure 4 that was constructed from the area-wide aggregate data. Conceptually these two sets of contribution estimates are not equivalent because we have not aggregated the country-level results (and because these are only four of the twelve euro area countries). As explained above, a major reason for cross-checking area-wide results with the country-level evidence was precisely to avoid the aggregation problems arising from the lack of a common monetary policy. Despite the underlying differences between the two distributions, we find them interesting in two respects. First, the contributions from these four major European countries do show investment contributions to be dominant. For instance, even at the twelve-quarter horizon, 60 percent of the consumption contributions are less than one-half. Second, the combined country-level distribution shows consistently lower consumption contributions than are found in the U.S. VARs.

Finally, the national structural model (point) estimates also confirm the larger role of investment contributions, except for France. It should be noted that the investment demand equation of the Banque de France model did not include a proxy for the user cost of capital at the time when this simulation exercise was performed.²² Thus, it is not surprising that the investment contribution in France according to this model was so low. While there are undoubtedly more subtleties to the country-level findings than our discussion suggests, it nonetheless

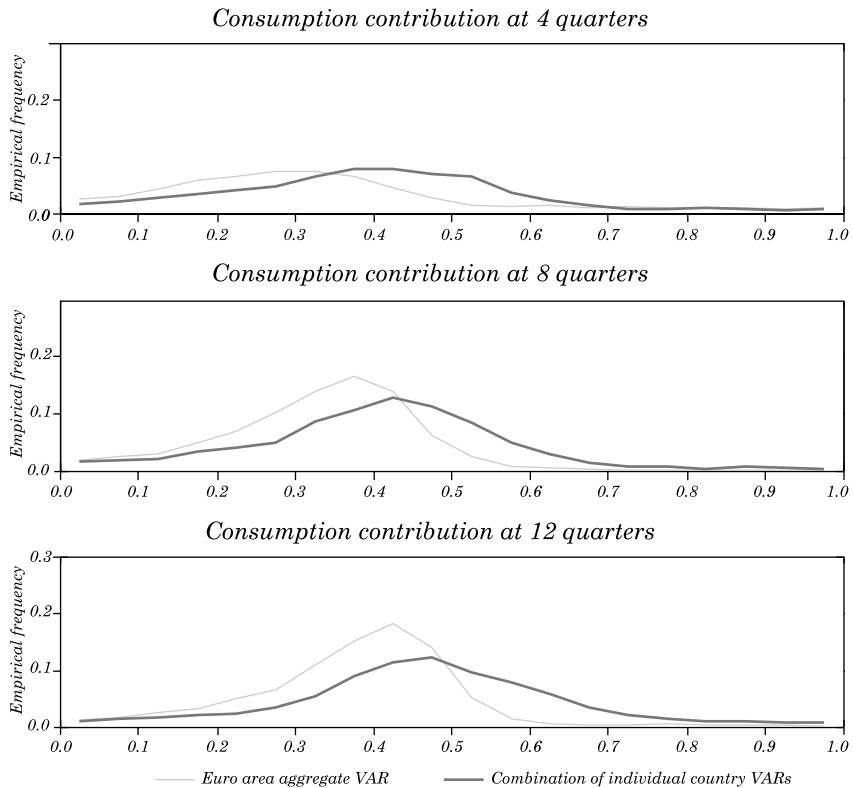
22. The large response of investment to monetary policy shock is also a feature of Mojon and Peersman (2003) estimations.

Table 6. Country-Level Evidence on the Contribution of Consumption to the Private Sector Domestic Demand Response to Policy Shocks
Percentiles and point estimates

Specification	Horizon	Country and percentile ^a											
		Germany			France			Italy			Spain		
		10th	50th	90th	10th	50th	90th	10th	50th	90th	10th	50th	90th
VARs	4 quarters	-1.92	0.64	5.16	-0.40	0.32	0.60	0.30	0.48	0.60	-0.40	0.32	0.90
	8 quarters	-0.38	0.36	1.80	-0.20	0.44	0.60	-0.20	0.36	0.60	0.10	0.39	0.60
	12 quarters	0.07	0.58	0.95	0.30	0.48	0.70	-1.00	0.43	1.90	0.20	0.42	0.70
NCB large-scale models (point estimates)	4 quarters		0.31			0.82			0.47			0.12	
	8 quarters		0.26			0.77			0.26			0.29	
	12 quarters		0.00			0.77			0.16			0.35	

a. Percentiles are based on 1,000 Monte Carlo simulations.

Figure 5. Distributions of the Consumption Contribution Calculated from the Euro Area Aggregate VAR and from the Combination of Individual Country VARs



Note: The density curves are based on 6,000 simulations for the euro area (1,000 draws for each of the two samples for each of the three VARs) and 4,000 simulations for the euro area countries (1,000 draws for each of the four countries).

seems safe to conclude that the country-level results and those for the area as a whole are broadly consistent. It appears that in both cases, the investment contributions are, relative to consumption contributions, substantially larger than in the United States.

A full investigation of the contributions across other economies is outside the scope of this paper. We note in passing that there are a couple of other pieces of evidence on this. We estimated a VAR for the United Kingdom and found that consumption contributions there were larger even than in the United States (results are available on

request). Conversely, Fujiwara (2003) estimates a set of VARs for Japan and finds that investment contributions there are much larger than consumption contributions. We look forward to further work aimed at establishing the output composition in other countries, but for the remainder of this paper we focus on the differences between the euro area and the United States.

3. INTERPRETING THE DIFFERENCES IN THE COMPOSITION OF OUTPUT EFFECTS

Our starting point is to check whether we can replicate the different compositions by appropriately choosing the parameters in small-scale dynamic stochastic general equilibrium (DSGE) models otherwise calibrated to fit the main features of the transmission mechanisms of the two economies. In this way, we should be able to trace the observed compositional differences back to a (hopefully small) set of differing structural features of the economies. These, in turn, could be checked against independent evidence, to arrive at a reasonably robust interpretation of our finding. Before embarking on this task, we quickly review the basic structure of this class of models.

3.1 DSGE Models in a Nutshell

Starting with the seminal works of Yun (1996), King and Wolman (1996), and Rotemberg and Woodford (1998), a growing body of literature has focused on extending the basic real business cycle (RBC) model to include a number of “real world” rigidities to account for some of the features of the data that the basic RBC model was unable to match. In this task, the main challenge was to remain firmly grounded in the optimizing behavior of a small set of rational, forward-looking representative agents (a consumer, a firm, possibly a financial intermediary, plus of course a government or a central bank) while incorporating a rich enough set of constraints limiting their decisions to fit the data. The constant challenge is to do all this while retaining numerical, if not analytical, tractability.

The challenge was met, with success, by skillfully combining four key ingredients. The first is a specification of the technology and of the market structure, originally due to Dixit and Stiglitz (1977). This assumption allows for product differentiation that is also compatible

with aggregation, so that overall economy-wide prices and quantities can be constructed.²³

The second critical ingredient is the assumption that prices and wages are set in the fashion proposed by Calvo (1983). This price- and wage-setting assumption, coupled with the assumed availability of a rich enough set of insurance markets, makes individual firms' prices (and wages) sticky, and this stickiness carries over to the aggregate levels of wages and price. One major advantage of this modeling strategy is that aggregate levels can be computed without having to keep track of all possible histories of previous pricing decisions.²⁴

The final two ingredients are a clever technique of log-linearization around a nonstochastic steady-state equilibrium and the use of efficient solution techniques for linear, rational expectation models. The (solved) theoretical model has then been matched with the data, combining calibration, matching of (selected) moments, or, more ambitiously, full maximum likelihood (cum Kalman filtering) estimation. Particularly good examples of what can be achieved along this route are, for the U.S., the model developed by Christiano, Eichenbaum, and Evans (2001) and, for the euro area, the model developed by Smets and Wouters (2002). Very recently the latter model has been estimated also for the United States (Smets and Wouters, 2003), and we use these estimates in what follows to try to develop an interpretation of differences in the composition of the output response.²⁵ We also mention some results obtained by Lindé (2003) with the Christiano, Eichenbaum, and Evans model.

These two models are indeed very similar. Relative to the first generation of DSGE models, they both embody a number of notable features aimed at improving the fit. First, together with the so-called Calvo adjustment for prices and wages, an assumption is made of full (in Christiano, Eichenbaum, and Evans) or partial (in Smets and Wouters) indexation to previous-period inflation for those agents that are not allowed to optimally reset their price (wage). This introduces

23. The aggregator is, however, of a constant elasticity of substitution (CES) nature, and it therefore differs from the linear aggregator that underlies national income and product accounts data.

24. In particular, it is the assumption that firms (households) can fully insure against the possibility of not being able to optimally set their price (wage) that makes that possibility a matter of irrelevance as far as the wealth of different agents is concerned, and therefore allows for a history-independent description of the economy's developments.

25. We are very grateful to Frank Smets and Raf Wouters for providing us with the model code.

inertia in the inflation process, a key feature of the data that a purely forward-looking formulation is unable to match.

Second, firms can optimally choose the intensity with which they use installed capital. Increasing (or decreasing) the utilization rate is not costless, and the firm balances the benefit of a marginal increase with its cost. Allowing capital services to be elastic and, in particular, to fall after a monetary policy tightening has the consequence of muting fluctuations in the (future) rental rate of capital, thereby helping to generate the gradual price response observed in the data; moreover, it also reduces the increase in labor productivity that would otherwise occur, thus offsetting the real effect of the tightening.

Third, consumers exhibit habit formation (in the Smets and Wouters model, the habit formation takes an “external form”, where the “habit” is provided by aggregate consumption, outside the control of the single household; in Christiano, Eichenbaum, and Evans, instead, the habit is proportional to the household’s own past consumption). This feature of the model is needed to get a gradual and hump-shaped response of consumption to a monetary impulse (observed in the data above). Indeed, the concavity of the utility function implies that a rise in the real interest rate (a fall in the price of future consumption, relative to present consumption), should be associated with low current consumption relative to the future, that is, with a counterfactually front-loaded response of consumption to the shock. Habit formation in essence makes the argument of the utility function to be (roughly) the *growth rate* of consumption, rather than its *level*. With this specification the hump-shaped response of consumption observed in the data after an interest rate increase is a consequence of the desire to make the growth rate low (more negative) today relative to tomorrow.

Fourth, changing the stock of capital (that is, investing) involves a cost (of course, above the price to be paid for the new machines). The role of the adjustment cost, much like the assumption of habit formation in consumption, is to prevent a front-loaded response of investment. In particular, any shock (including the types of monetary policy ones considered above) that generates persistent changes in real interest rates will engender (absent adjustment costs) a substantial and immediate drop in investment. Adjustment costs, modeled as penalizing the change in investment, prevent this counterfactually large and immediate response.

While these four features do not exhaust the richness of the two models, they are arguably what enables them to match many features

of the empirical transmission mechanism much better than plain-vanilla RBC models do.

It is probably too early to judge whether these models, and DSGE models more generally, will live up to the challenge of replacing the more traditional large-scale econometric models in use by many decisionmakers and practitioners. DSGE models certainly have a number of advantages, notably delivering a set of rigorously grounded theoretical and econometric findings that still adequately fit the data. However, these models have some limitations that might complicate their use in trying to explain our puzzle. In particular, the DSGE models typically assume the availability of a complete set of markets, thus making it difficult to generate precautionary behavior or liquidity constraints that might affect the consumption response to changes in the monetary policy stance. Related to this, we don't know of a DSGE that can explain the different historical patterns of financial market developments. This is one of the most striking differences between the United States and the euro area that might have a bearing on the observed differences in the composition of the output response to monetary policy. It is also worth recalling that the representative-agent nature of these models makes them liable to potential pitfalls resulting from aggregation problems (see Kirman, 1992; Altissimo, Siviero, and Terlizzese, 2002), whose actual importance still needs to be assessed.

Nonetheless, we believe these models are rich enough to provide us with an organized way to interpret the evidence. In particular, they have a number of features—pertaining both to short-run frictions and to long-run equilibrium properties—that make them potentially suitable for identifying the determinants of the puzzle.

3.2 Examining the Output Composition in the Smets-Wouters Model

We use the Smets-Wouters model to address three questions. First, are there identifiable mechanisms that are, in principle, capable of generating a difference in the output composition in the United States and euro area? Second, do the different estimates of the parameters of the model for the two economies imply a difference in the output composition that is *qualitatively* similar to the section 2 findings? Third, are these implied differences in the output composition *quantitatively* in line with the above evidence?

To our first question the answer is a clear yes, as there are several features of the model that could generate different output compositions. Our experimentation simulating the model suggests, however, that not all of the mechanisms present in the Smets-Wouters model matter for the composition of the output response. We determined that out of the large number of parameters estimated, only five parameters make any material difference for the implied consumption contributions (in the wake of a shock to short-term interest rates) at horizons up to twelve quarters after the initial shock.

Four of these five parameters govern fairly intuitive economic mechanisms. The first of these determines the size of investment adjustment costs. Higher values of this parameter dampen investment responses and hence yield relatively larger consumption contributions. In light of these adjustment costs, transitory changes in user cost of capital will have limited effects. This means that the parameter in the central bank's reaction function that measures the amount of interest rate smoothing is also important. The more the persistence that the central bank induces in the (real) short-term interest rate, the larger are the predicted investment responses.²⁶ A third key parameter measures the strength of the habit persistence in consumption. When habit persistence is stronger, then the adjustment in consumption following an interest rate shock is smaller. Finally, the level of the intertemporal elasticity of substitution is also relevant. When consumers have high willingness to shift consumption over time, a given interest rate change will engender larger consumption responses.

In addition to these behavioral parameters, the share of capital in the (long-run) Cobb-Douglas production function matters. The mechanisms associated with this parameter are less intuitive and somewhat more mechanical. The stability of the model requires the capital-to-output ratio to be restored (eventually) after all shocks. This means that the long-run movement in investment must move in proportion to the long-run change in output. The Cobb-Douglas parameter therefore has two influences on the contributions. First, it plays

26. This is the one exception to our earlier claim about the robustness of the contribution statistics, since changes in this persistence parameter can correspond to very different experiments that need not be comparable. The intuition given above that presumes that persistence increases the importance of investment (implicitly relative to consumption) depends on investment being more interest sensitive than consumption. This is true for the baseline Smets-Wouters parameter estimates.

a role in the initial steady-state level of capital to output that is in place at the start of any simulation. Second, in the wake of any shock, the shape of the investment response is left open, but the total amount of adjustment is constrained by the Cobb-Douglas parameter. In particular, for an economy with a relatively high capital-to-output ratio, more investment adjustment must occur for any given change in the level of output.

Given that the model can, in principle, generate different output compositions, we next ask whether this, in fact, occurs for the (modal) point estimates presented in Smets and Wouters.²⁷ Our answer is a qualified yes. It turns out that the implied consumption contributions do differ in a way that qualitatively matches the patterns found in the data. However, this does not appear to result from the more intuitive, behavioral channels discussed above, and, most importantly, it hinges on parameter differences that are not estimated.

The consumption contributions implied by the baseline estimates for the United States are, at quarters four, eight, and twelve, respectively, equal to 0.54, 0.48, and 0.45. Those in the euro area are instead 0.48, 0.40, and 0.36.

The mechanisms that drive this result do not depend on the intuitive channels because they tend to offset each other. The intertemporal elasticity of substitution is estimated to be lower in the euro area, and the persistence in the monetary policy reaction function higher; both features imply relatively lower consumption contributions there (as it is in the data). Conversely, the adjustment cost parameter is estimated to be larger in the euro area, and the habit formation parameter smaller; both features imply relatively higher consumption contributions (contrary to what observed in the data). Besides working in opposite directions, the quantitative differences in these four parameters are modest: none of these four parameters is estimated to be much more than one standard deviation different between the two economies.

While these channels largely cancel, the ones tied to the Cobb-Douglas parameter do not. In the Smets-Wouters model this parameter is not estimated, however, but instead is fixed so that it, along with the subjective discount rate and depreciation rate (assumed equal in the two economies), implies a steady-state value of the ratio of investment to output that matches its observed sample average. To this end, the value of the capital share in the euro area is set to be

27. Since their estimation procedure is not explicitly geared to reproducing this specific dimension of the data, this need not be the case.

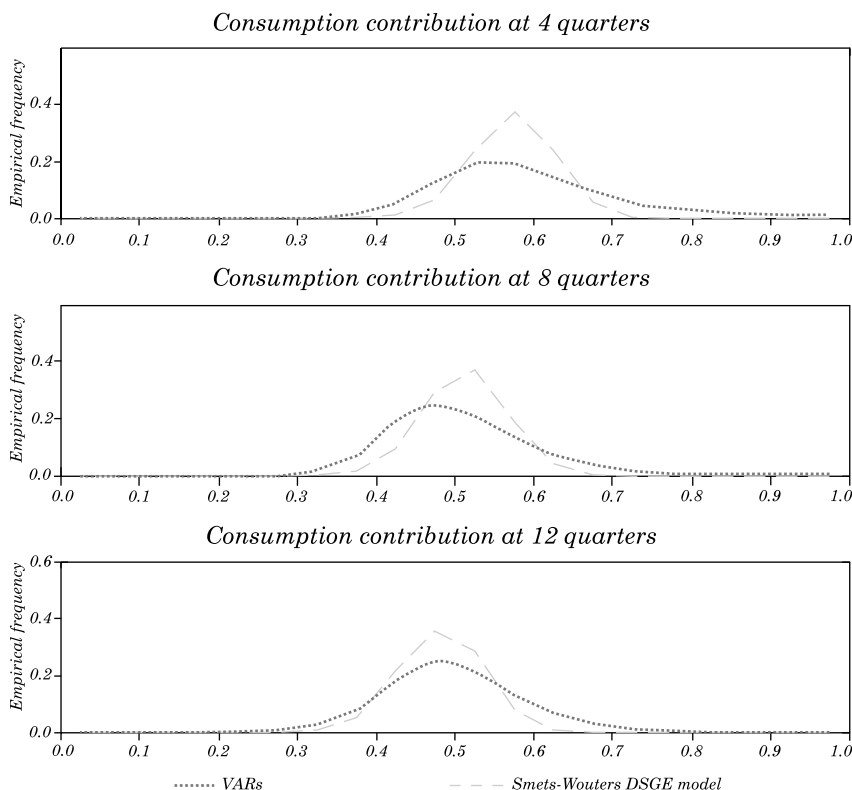
higher (0.30 versus 0.24 in the United States). The model is thus hard-wired to deliver larger long-run investment responses in the euro area. This constraint is, however, unlikely to be relevant within the twelve-quarter horizon that we consider: the long-run responses only tend to prevail much later, between ten and fifteen years after the initial shock. But the short-run values of the contributions are proportional to the starting capital to output levels. Thus, if the semi-elasticity of investment to the interest rates is similar in the two economies, then the amount of overall investment (and the investment contributions we compute) will be higher in the euro area. It turns out that the semi-elasticities generated by the model are indeed similar—at least up to quarter eight—but because of different assumed initial conditions these similarities lead to bigger investment contributions and lower consumption contributions in the euro area.

We now turn to our last question, of whether there is quantitative consistency between the model generated and the observed differences in the output composition. Our answer is, not really. The consumption contributions implied by the point estimates mentioned above already show that the differences, though of the right sign, are not nearly as large as those documented in section 2. To arrive at a more systematic and robust assessment, we conducted a Monte Carlo exercise identical to the earlier ones done for the VARs: this is done by drawing the full set of estimated parameters 1,000 times from the posterior distribution and computing the implied consumption contributions at the four-, eight-, and twelve-quarter horizons. We also reestimated the VARs over the 1974 to 2001 period (because the Smets-Wouters parameters were estimated over this period). The resulting distributions for the Smets-Wouters model and the VARs are shown in figures 6 (for the euro area) and 7 (for the United States).²⁸

Figure 7 shows that drawing from the estimated distribution of the parameters for the United States, the Smets-Wouters model generates a distribution of contributions that is relatively close to that based on the VARs, at least at a twelve-quarter horizon. Figure 6 shows that a similar experiment for the euro area yields distributions rather different from those based on VARs—the Kolmogorov-Smirnov tests (at all three horizons) reject the equality of these distributions at significance levels below 1 percent.

28. The median consumption contributions of the model generated distributions in the United States are, at quarters 4, 8 and 12, respectively, equal to 0.57, 0.51, and 0.48, those in the euro area are 0.50, 0.42, and 0.37. These values are close to those reported above in the text, based on the point estimates.

Figure 7. Distributions of the Consumption Contribution in the U.S., VARs and Smets-Wouters DSGE model



Note: The density curves are based on 3,000 simulations for the VARs (1,000 draws for each of the 3 VARs) and 1,000 simulations for Smets and Wouters model (1,000 draws from the joint distribution of the estimated model parameters).

observed in the data. Depending on whether one compares the Smets-Wouters model to the large-scale models or the VARs, it is possible to get the Smets-Wouters model to work for one economy or the other, but not both.

A different way to see this challenge is to ask how big the estimated parameter differences would have to be in order to generate substantial differences in the implied contributions coming from the model. We need relatively large changes in any of the relevant parameters to generate contributions that are close enough to what is

observed in the data. For example, if we want the model to generate point estimates of the contribution that are roughly equal to the median of the VAR distribution for the euro area (something around 0.25 at each of the three horizons), we need to reduce (relative to the model's baseline) the adjustment cost parameter by more than four times its estimated standard error, or the intertemporal elasticity of substitution by almost four times.²⁹ Alternatively, we need to increase the habit formation parameter by about six times its standard error, or the persistence in the policy reaction function by about five times.³⁰

It is interesting to note that Lindé (2003), estimating the Christiano, Eichenbaum, and Evans model for both the euro area and the United States, seems to be able to account for the different patterns in the output composition. The estimation procedure adopted by Lindé, in line with the approach originally followed by Christiano, Eichenbaum, and Evans (2001), is less ambitious than that pursued by Smets and Wouters, as it simply tries to match the impulse responses to a monetary policy shock. Therefore, since the output composition that we are trying to reproduce is a feature of those impulse responses, Lindé's result is not really surprising.

It should be noted, however, that the changes in the parameter estimates obtained by Lindé are quite large. In particular, the parameter capturing adjustment costs in investment is, according to Lindé's estimates, thirteen times smaller in the euro area relative to the United States. (There are also other big changes, but we believe they are not really important for the ability of the model to reproduce the output composition pattern.) We see these differences as implausibly large.

Summing up, we conclude that the mechanisms at play in the most recent generation of DSGE models that might potentially account for the output composition puzzle (adjustment costs in investment decisions, habit formation, interest rate smoothing, or willingness to substitute present for future consumption) do not provide a fully satisfactory explanation. Therefore, we believe it is appropriate to explore further the set of potential explanations of the output composition puzzle, relaxing the constraints posed by this version of the models.

29. Both of the mentioned changes generate about the right contribution at quarters 8 and 12, but yield too large a contribution at quarter 4.

30. Even this change, which makes autoregressive component in the policy rate very close to 1, is not enough to match the observed contributions.

3.3 Are the Differences Due to Consumption or Investment?

Our analysis of section 2, was based on “contributions” precisely because we viewed these measures as being relatively robust across models and definitions of monetary policy shocks. A drawback of the contribution statistic is that because it is a ratio, it does not allow us to identify whether the consumption contributions are higher in the United States because U.S. consumers respond more than euro area consumers to a monetary policy shock or because U.S. firms vary their investment less. To identify this we need to look at the *levels* of the responses. These are difficult to compare. Nevertheless, several pieces of evidence suggest that that consumption is more likely to be at the root of our puzzle. Next, we offer some preliminary conjectures as to why consumption might be more responsive to a monetary policy shock in the United States.

We start by examining the structural model simulations. In these models, one can easily trace the effects of the same exogenous interest rate (and exchange rate) path on all the variables in the model, including consumption and investment, and compare the results. While this experiment has the weakness of suspending the policy reaction functions, at least it allows for a neat comparison. These simulations suggest that investment responses are surprisingly similar. In the FRB/US model, the drop is about 0.3 percent relative to the baseline value in the first year, about 1.8 percent in the second year, and about 3.1 percent in the third. In the euro area models, the drop is in the range 0.3 percent to 0.8 percent in the first year, 1 percent to 2.4 percent in the second, and 1.2 percent to 3.0 percent in the third (see table 3).

In contrast, there appear to be large differences in the response of consumption to the policy rate shifts. In the FRB/US model, the drop is about 0.4 percent of the baseline value in the first year and about 1.4 percent in both the second and third years; in the euro area models the drop is in the range 0.1 percent to 0.3 percent in the first year, 0.2 percent to 0.6 percent in the second, and 0.2 percent to 0.5 percent in the third.

Turning next to the VARs, the estimated profile of the investment response to monetary shocks is rather similar in both areas, with the drop peaking about one and half years after the shock and a gradual return to baseline afterwards. Once the differences in the size of the initial shock are broadly taken into account, the magnitude of the

(maximum) drop is also roughly similar. For example, for the two Christiano, Eichenbaum, and Evans (2001) specifications (using the longest samples), the maximum drop is slightly bigger than 1 percent in the United States, after an initial interest rate shock equal to 0.7, as compared to about 0.75 percent, following an initial interest rate shock equal to 0.4 for the euro area version.

The VARs also seem to show quantitative differences in consumption responses to a monetary shock. The point estimates for all three euro area models display a mild hump-shaped pattern, but the standard errors suggest that the responses are typically not different from zero. The 90 percent confidence intervals in the long sample typically suggest that the biggest response would be no more than -0.3 . For the United States, the consumption responses are significantly different from zero, and the point estimates for the peak responses for all three models in the long sample exceed -0.25 .

Overall, we read this evidence as casting doubt on an explanation based on differences in the investment response, instead pointing to consumption differences as the most likely culprit. We therefore proceed by considering explanations for why consumption in the two economies might respond differently to a monetary policy shock.

3.4 Alternative Explanations for the Consumption Differences

One class of explanations that appears intuitively appealing focuses on the relative degrees of social insurance in the euro area versus the United States. In particular, the availability of complete insurance that is assumed in the DSGE model that we have examined limits the extent to which issues pertaining to, for example, precautionary saving in the face of employment or labor income risk can be explored. Yet most of the literature suggests that this is a source of major differences in the structure of personal incentives in the two economies. It is natural to think that this element should affect consumer behavior, as it is believed to affect, for example, labor supply. Hence, we see mechanisms that involve insurance market incompleteness as a natural avenue for exploration, and we focus on this in the final section of the paper.

We examine a select number of potential mechanisms that could give rise to the observed differences. Our first candidate focuses on potential effects of labor market risk in the two economies, focusing specifically on unemployment. The permanent income theory of

consumption suggests focusing on the (average) cumulative effect of the shock, so that besides the short-run or peak effects, the duration of any labor market adjustments will matter. This complicates the task, because we believe most prior work suggests that there is more (unconditional) short-run unemployment risk in the United States, but that the risk of a long spell is larger in Europe.³¹ We proceed by examining the mean effects of unemployment of monetary shocks as implied by both the VARs and the structural models.

The results obtained from the central banks structural models support the idea that unemployment spells are more likely to result from a monetary policy shock in the United States. In the FRB/US, unemployment increases by 0.12 percentage points in the first year, by 0.56 in the second, and by 0.77 in the third; in the euro area models, the increase is in the range of 0.04 to 0.08 percentage points in the first year, 0.11 to 0.36 in the second, and 0.17 to 0.61 in the third (see table 3).

The VAR evidence suggests otherwise. This evidence is compiled by adding unemployment to the models we used earlier as an extra variable, without changing the identification procedure for the monetary shocks. Table 7 shows the results for the Erceg and Levin model and the Peersman-Smets model without M3 for the long sample periods—the results for other models and sample periods are similar. For both VARs, unemployment is hardly estimated to change in response to a monetary shock (with effect being less than 0.1 percentage points at all horizons for both models).

Given the conflicting findings from the two methodologies, the explanation of the puzzle based on differences in labor market risk in the two economies cannot be confirmed. The issue remains open until further evidence is available. Meanwhile, we explore other possible alternatives.

A closely related possibility is that the combination of more generous unemployment benefits, national health care systems, and generous pay-as-you-go pension schemes all help to insulate euro area residents more from adverse economy-wide shocks than Americans. For example, Martin (1996) compares benefit replacement rates (net of housing and taxes) for households of varying family size across the United States and European countries. Roughly speaking, the replacement rates in most European countries are at least twice as high in the first year of unemployment as in the United States and five to ten times higher in subsequent years. The latter result is due to the U.S.

31. For instance, Bean (1994) shows that transitions probabilities both in and out of unemployment are larger in the United States than in Europe.

Table 7. Effects of Monetary Policy Shocks on Disposable Income and Unemployment^a

Deviation from baseline

Variable	Horizon	Region and sample				
		United States			Euro area	
		1960-2001	1965-1979+ 1984-2001	1984-2001	1980-2000	1970-2000
Disposable income	4 quarters	-0.15	-0.15	-0.02	-0.02	-0.02
	8 quarters	-0.16	-0.04	-0.14	-0.15	-0.17
	12 quarters	-0.07	0.08	-0.21	-0.10	-0.18
Cumulated disposable income	4 quarters	-0.33	-0.28	-0.02	0.14	0.04
	8 quarters	-0.93	-0.53	-0.39	-0.32	-0.46
	12 quarters	-1.33	-0.38	-1.16	-0.84	-1.21
Unemployment	4 quarters	0.05	0.05	0.05	0.04	0.03
	8 quarters	0.07	0.08	0.14	0.09	0.09
	12 quarters	0.04	0.02	0.10	0.07	0.06

a. The U.S. model is the Erceg–Levin VAR, while the euro area model is the Peersman–Smets model without M3. Authors' calculation of the impulse response function of the variables of interest in VARs presented in section 1 of the paper. Disposable income and unemployment were included after the three GDP components in the baseline VARs.

benefits dropping sharply after the first year, so that most families could not expect to recover even 15 percent of their income.

However, while evidence suggests that *social* protection is higher in the euro area, *market* protection, either in the form of straight insurance markets or in the form of risk-sharing transfers taking place among regions, is arguably bigger in the United States. In fact, the typical finding from the literature on risk sharing (for example, Asdrubali, Sorensen, and Yosha, 1996; Sorensen and Yosha, 1998; Forni and Reichlin, 1999) is that it is low in European countries compared with the United States.³²

On our reading, the main focus in this literature is not on the overall amount of uninsured risk that consumers in the two areas ultimately end up bearing—which is what matters for our purposes—but rather on the forms and relative importance of implicit insurance mechanisms. One result, from Forni and Reichlin (1999), suggests that the variance of income is higher in the United States at business

32. For instance, Forni and Reichlin (1999) write that “the extent of risk-sharing through capital markets and EC structural funds is very little if comparison is made with the United States.”

cycle frequencies and is higher in Europe at long-run frequencies. We are unaware of any direct studies that focus on risk sharing in the wake of monetary policy shocks per se.

As an admittedly indirect attempt to assess this, we look at the effect of the policy shock on aggregate personal disposable income, both in central banks structural models and in VARs. Looking at central banks models, the evidence is mixed. Comparing the FRB/US model with the aggregation of NCBs models (see table 3), the response of disposable income to a monetary policy shift is about the same up to the first year, but is substantially smaller (by something between three and five times) in the euro area in the following two years. If the comparison is made with the area-wide model (AWM), however, the response of disposable income is larger in the euro area in the first year and of the same order of magnitude in the following two years.

The associated VAR evidence is shown in table 7. We again show the results for the Erceg and Levin specification for the United States and the Peersman and Smets model without money for the euro area (but the results from the other models are similar). We find some evidence that the response of disposable income is somewhat stronger and quicker in the United States. In particular, for the first few quarters, disposable income increases in the euro area, while it falls immediately in the United States. In terms of the cumulated response of disposable income in the two areas, the response in the United States is substantially larger for the first two years. However, the gap seems to close during the third year.

While the overall findings are somewhat ambiguous, it is striking that across models the relative movements in consumption and disposable income in both economies following a monetary policy shift are similar. In the euro area both the VARs and the structural models suggest that disposable income and consumption move roughly in proportion to each other, whereas the FRB/US model and the U.S. VARs suggest that the decline in consumption is roughly twice the size of the decline in disposable income.

We view this finding as certainly meriting further exploration. For instance, marshalling all available data on the components of the budget and nature of transfer programs and checking whether these transfers effectively buffer the risk associated with a given decline in disposable income would be an interesting next step.³³

33. Potentially larger response of consumption relative to disposable income could be due to capital market imperfections, if one believed that U.S. households faced more severe borrowing constraints, something that we believe is doubtful.

A final possibility is that the differences in the consumption response could be due to different wealth effects of monetary policy. Reliable comparable data on the structure of wealth is hard to obtain, but it is widely believed that the structure of wealth holding differs across the two economies. For instance, total financial assets in the hands of households were, in 2001, 321 percent of GDP in the United States and 202 percent in the euro area (Agresti and Claessens, 2003). Besides the tendency of Americans to hold more of their wealth in financial securities, the form of the holdings appears to differ. The Europeans tend to hold more government debt in their portfolios than Americans, whereas Americans hold relatively more equity market claims and corporate debt. But given the limited detail available on the holdings, we cannot directly measure the response of most components of wealth to changes in interest rates. We thus cannot assess this channel.

4. CONCLUSIONS

Our focus in this paper is a comparison of certain key macroeconomic features of the transmission mechanisms of monetary policy between the United States and the euro area. After the establishment of the euro area as the largest currency area in the world, with a new and independent central bank, a comparative understanding of the two transmission mechanisms has, in our view, become important. Looking at them together can not only sharpen our understanding of each and identify clues as to where and why they differ, but also allow us to better appreciate the global implications that the independent conduct of monetary policy in each of the two areas generates.

We proceed in steps. We first compare the cyclical properties of euro area and U.S. macroeconomic time series. Here the striking fact, already reported by other recent papers, is that such properties are, in fact, broadly similar, suggesting that common underlying market forces are at work.

Next we analyze a small set of VAR models for the two areas. We find that, again, the main macroeconomic facts are similar. Specifically, after a monetary shock, real GDP displays a hump-shaped profile, returning to baseline, whereas the price level diverges gradually but permanently from the initial value. Thus, the consensus on the way monetary policy operates in the United States has held up through the long business cycle expansion of the 1990s. Moreover, the consensus view seems to well describe the euro area facts, too.

However, prior work has paid relatively little attention to the underlying adjustments that accompany the change in output. In this respect the two areas differ. In particular, after a change in monetary policy, the role of household consumption in driving output changes is greater, and that of investment smaller, in the United States relative to the euro area. This difference is present in VAR estimates and those of large-scale structural econometric models. We call this the output composition puzzle.

To explore and explain the puzzle we take two tacks. First, we consider a class of dynamic stochastic general equilibrium models. Our main result here is that these models, at least in the versions that are now considered on the research frontier, have difficulty fully accounting for the puzzle. The (full information) estimates of one such model do not quite deliver the pattern present in the data.

Given this conclusion, we turn to several less tightly structured tests and hypotheses. We first make a tentative assessment of whether the puzzle is more likely due to divergent behavior of consumers or firms. It appears to us that the consumers are responsible for the differences. Unfortunately we do not have a compelling explanation for why this is the case. It appears that disposable income may be less responsive to monetary changes in the euro area than in the United States. We were motivated to make this comparison by the hypothesis that the social safety net in Europe might cushion the effects of monetary policy on consumption more there. It appears that movements in consumption relative to disposable income are larger in the United States, too. Explaining this finding and sharpening the tests of the conjecture about the importance of the social safety net are an obvious next step.

APPENDIX A

Country-level VARs

Country level VARs were estimated for France, Germany, Italy, and Spain, for which we could gather, thanks to Alberto Locarno, unpublished private sector investment series. Except for the inclusion of consumption and investment inside the VARs, these country-level VARs are quite similar to the ones presented in Mojon and Peersman (2003).

In Germany, the VAR model includes consumption, investment, what we call the rest of GDP, inflation, the short-term interest rate, the real effective exchange rate, a trend, and a constant. For the other countries, we include the same variables plus the German interest rate in order to account for the anchoring role of the Bundesbank monetary policy during the EMS. The models are estimated for the period following the launch of the EMS (1980 to 2001) and the identification of the French, Italian, and Spanish monetary policy shocks are performed by a Cholesky decomposition, ordering the domestic money market interest rate last.

In the case of Germany, we solve the endogeneity bias between the interest rate and the exchange rate innovations by imposing that a 1 percent interest rate shock triggers a simultaneous 1 percent appreciation of the effective exchange rate. While arbitrary, this identification assumption solves the price puzzle with a smaller appreciation than the one obtained by instrumental variable estimation in Mojon and Peersman (2003).

There are three other differences with the country-level VARs of Mojon and Peersman (2003). First, we use the European System of Accounts 1995 (ESA95) national account data and we extend the sample period by three years, from 1999 to 2001.³⁴ Second, we choose the effective exchange rate variable rather than the bilateral rate to the deutsche mark for France and for Spain. This seems more appropriate given that the bilateral rates converged to their final parity in the second half of the 1990s and then remained constant after the launch of the euro. This change of variable was, however, not feasible in the case of Italy. There, we kept the lira–deutsche mark exchange rate, and we included the same set of exogenous variables as that used in the PS model of the euro area in order to obtain well-behaved responses

34. In the case of Germany, the data before 1991 were obtained by back dating the growth rates of the ESA79 West Germany national account data.

to interest rate shocks, that is, to avoid a persistent price and output puzzle. Finally, we estimated the German VAR with dummies that exclude nine “reunification” observations, from 1990:1 to 1992:1. These dummies allow us to eliminate the effect of the reunification period when interest rates, in parallel prices and GDP, increased in a somewhat unusual way. Mojon and Peersman (2003) managed to minimize the weight of this episode by estimating their German VAR on a sample that also included the 1970s. Having the estimates for the four countries for a sample for which we could compare the effect of monetary policy shocks with the outcome of the VAR estimated with synthetic euro area data seemed to us more appropriate.

APPENDIX B Data Sources

Euro Area

Most euro area time series are taken from the ECB Area-Wide Model database. These data are presented on page 51 in Fagan, Henry, and Mestre (2001). Updates of these series up to 2000:4 can be obtained from Alistair.Dieppe@ecb.int.

We use both the previously available historical time series for M3 (February 1999 monthly bulletin of the ECB) to conform with Peersman and Smets (2003) and the more recent series backdated to 1970 (not yet published) for the VAR models estimated over a sample covering the 1970s.

The stock price, available only from 1973 onward, is the EMU-DS market index of euro area stocks published by Datastream.

Aside from the historical M3 series dating to the 1970s and the HICP, all the series we use were already seasonally adjusted. We adjusted these remaining two series using the seasonal adjustment routine in Eviews.

Country-level Data

We use ESA95 national account data for GDP and consumption. Private investment series are obtained by subtracting public investment series from the total investment series that are available in ESA95. The public investment series come from the quarterly national account published by the statistics institutes in the case of Germany and France. The Italian and Spanish public investment series come from Banca d'Italia and Banco de España. We are grateful to Alberto Locarno for providing us with these series for the four countries.

In the case of Germany, national account series prior to 1991 were backdated using the growth rate of the West German ESA79 series.

The interest rates are the three-month money market interest rates, which, from 1999:1, were set equal to the euro area three-month money market rate. For France, Germany and Spain, we used the CPI based real effective exchange rate produced by the BIS, while for Italy, we used the lira-DM exchange rate (from the ECB's internal database).

United States

Most U.S. macroeconomic time series are downloaded from www.freelunch.com. We list the original source for the different series in table B1.

The private consumption series available from the Bureau of Economic Analysis starts only in 1967. To arrive at a longer time series, we added the nondurable goods, durable goods, and services consumption series provided to us by Larry Christiano. He also supplied us with the real wage and labor productivity data that we use. These series were downloaded from economics.dri-efa.com/webstract.

Finally, the profits series corresponds to the corporate after tax profits as available in the Bank for International Settlements (BIS) database.

Table B1. Sources of U.S. Macroeconomic Time Series

<i>Series</i>	<i>Source</i>
GDP and all GDP components	Bureau of Economic Analysis
CPI: urban consumer—all items, (1982–84 = 100, SA)	Bureau of Labor Statistics
Commodity price index	KR–CRB futures price index, (1967 = 100), Knight–Ridder
Stock price index 500 composite	Standard & Poor's, (index 1941–43 = 10, month end)
Federal funds rate	Federal Reserve Board: H.15
Ten-year constant maturity securities	Federal Reserve Board: H.15
Total reserves and nonborrowed reserves (adjusted for changes in reserve requirements, millions \$, SA)	Federal Reserve Board: Aggregate reserves of depository institutions—H.3
M1 and M2, (SA billions \$)	Federal Reserve Board: H.6 Money stock and liquid assets, and debt measures

REFERENCES

- Agresti, A.M. and J. Claessens. 2003. "The Euro Area Economic and Financial Structure: An Overview." In *Monetary Policy Transmission in the Euro Area*, edited by I. Angeloni, A. Kashyap, and B. Mojon. Cambridge University Press: 433–54.
- Agresti, A.M. and B. Mojon. 2001. "Some Stylised Facts on the Euro Area Business Cycle." Working paper 95. Frankfurt: European Central Bank.
- . 2003. "Some Stylised Facts on the Euro Area Business Cycle." In *Monetary Policy Transmission in the Euro Area*, edited by I. Angeloni, A. Kashyap, and B. Mojon. Cambridge University Press: 15–35.
- Ahmed, S., A. Levin, and B.A. Wilson. 2002. "Recent U.S. Macroeconomic Stability: Good Luck, Good Policies, or Good Practices?" Working paper 730. Board of Governors of the Federal Reserve System, International Finance Division.
- Altissimo, F., S. Siviero, and D. Terlizzese. 2002. "How Deep Are the Deep Parameters?" *Annales d'Economie et Statistique* 67–68: 207–26.
- Angeloni, I., A. Kashyap, and B. Mojon, eds. 2003. *Monetary Policy Transmission in the Euro Area*. Cambridge University Press.
- Asdrubali, P., B.E. Sorensen, and O. Yosha. 1996. "Channels of Interstate Risk Sharing: United States, 1963–90." *Quarterly Journal of Economics* 111(4): 1081–110.
- Barth, M.J. and V.A. Ramey. 2002. "The Cost Channel of Monetary Transmission." In *NBER Macroeconomics Annual 2001*, edited by B.S. Bernanke and K. Rogoff, 199–239. MIT Press.
- Baxter, M. and R.G. King. 1999. "Measuring Business Cycles: Approximate Band-Pass Filters for Macroeconomic Time Series." *Review of Economics and Statistics* 81(4): 575–93.
- Bean, C. 1994. "European Unemployment: A Survey." *Journal of Economic Literature* 32(2): 573–619.
- Bean, C., J. Larsen, and K. Nikolov. 2003. "Financial Frictions and the Monetary Transmission Mechanism: Theory, Evidence and Policy Implications." In *Monetary Policy Transmission in the Euro Area*, edited by I. Angeloni, A. Kashyap, and B. Mojon. Cambridge University Press: 107–30.
- Bernanke, B.S. and M. Gertler. 1995. "Inside the Black Box: The Credit Channel of Monetary Policy Transmission." *Journal of Economic Perspectives* 9(4): 27–48.

- Bernanke, B.S. and I. Mihov. 1998. "Measuring Monetary Policy." *Quarterly Journal of Economics* 113(3): 869–902.
- Boivin, J. and M.P. Giannoni. 2002. "Has Monetary Policy Become More Effective?" Columbia University. Mimeographed.
- Calvo, G. 1983. "Staggered Prices in a Utility-maximizing Framework." *Journal of Monetary Economics* 12: 383–98.
- Christiano, L.J. and M. Eichenbaum. 1995. "Liquidity Effects, Monetary Policy and the Business Cycle." *Journal of Money, Credit and Banking* 27(4, part 1): 1113–36.
- Christiano, L.J., M. Eichenbaum, and C. Evans. 1999. "Monetary Policy Shocks: What Have We Learned and to What End?" In *Handbook of Macroeconomics*, vol. 1A, edited by J.B. Taylor and M. Woodford, 65–148. Amsterdam: North-Holland.
- . 2001. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." Working paper 8403. Cambridge, Mass.: National Bureau of Economic Research.
- Clarida, R., J. Galí, and M. Gertler. 2000. "Monetary Policy Rules and Macroeconomic Stability." *Quarterly Journal of Economics* 115 (February): 147–80.
- Dieppe, A. and J. Henry. 2002. "The Euro Area Viewed as a Single Economy: How Does It Respond to Shocks?" Frankfurt: European Central Bank. Mimeographed.
- Dixit, A. and J. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *Microeconomic Theories of Imperfect Competition: Old Problems and New Perspectives*. Elgar Reference Collection, *International Library of Critical Writings in Economics*, vol. 102.
- Erceg, C. and A. Levin. 2002. "Optimal Monetary Policy with Durable and Non-durable Goods." Working paper 179. Frankfurt: European Central Bank.
- Fagan, G., J. Henry, and R. Mestre. 2001. "An Area-Wide Model (AWM) for the Euro Area." Working paper 42. Frankfurt: European Central Bank.
- Forni, M. and L. Reichlin. 1999. "Risk and Potential Insurance in Europe." *European Economic Review* 43(7): 1237–56.
- Fujiwara, I. 2003. "Output Composition of Monetary Policy Transmission in Japan." Discussion Papers in Economics and Business 03-07. Osaka University, Graduate School of Economics and Osaka School of International Public Policy.
- Gordon, D. and E. Leeper. 1994. "The Dynamic Impacts of Monetary Policy: An Exercise in Tentative Identification." *Journal of Political Economy* 102(6): 1228–47.

- Kahn, J., M.M. McConnell, and G. Pérez-Quirós. 2002. "One of the Causes of the Increased Stability of the U.S. Economy." *Federal Reserve Bank of New York Economic Policy Review* 8 (May): 183–202.
- King, R. and B. Wolman. 1996. *Federal Reserve Bank of St. Louis Review* 78(3): 83–107.
- Kirman, A. 1992. "Whom or What Does the Representative Individual Represent?" *Journal of Economic Perspectives* 6(2): 117–36.
- Lindé, J. 2003. "Comment on the Output Composition Puzzle: A Difference in the Monetary Transmission Mechanism in the Euro Area and the U.S.." Stockholm: Sveriges Riksbank. Mimeographed.
- Ludvigson, S., C. Steindel, and M. Lettau. 2002. "Monetary Policy Transmission through the Consumption-Wealth Channel." *Federal Reserve Bank of New York Economic Policy Review* 8(1): 117–34.
- Mankiw, N.G. 2001. "The Inexorable and Mysterious Tradeoff between Inflation and Unemployment." *Economic Journal* 111(471): 45–61.
- Martin, J.P. 1996. "Measures of Replacement Rates for the Purposes of International Comparison: A Note." *OECD Economic Studies* 26: 99–115.
- Meltzer, A. 1995. "Monetary, Credit and (Other) Transmission Processes: A Monetarist Perspective." *Journal of Economic Perspectives* 9(4): 49–72.
- Mojon, B. and G. Peersman. 2003. "A VAR Description of the Effects of Monetary Policy in the Individual Countries of the Euro Area." In *Monetary Policy Transmission in the Euro Area*, edited by I. Angeloni, A. Kashyap, and B. Mojon. *Cambridge University Press*: 56–74.
- Obsfeld, M. and K. Rogoff. 1995. "The Mirage of Fixed Exchange Rates." *Journal of Economic Perspectives* 9(4): 73–96.
- Peersman, G. and F. Smets. 2003. "The Monetary Transmission Mechanism in the Euro Area: More Evidence from VAR Analysis." In *Monetary Policy Transmission in the Euro Area*, edited by I. Angeloni, A. Kashyap, and B. Mojon. *Cambridge University Press*: 36–55.
- Ramey, V. and D. Vine. 2003. "Tracking the Source of the Decline in GDP Volatility: An Analysis of the Automobile Industry." University of California at San Diego. Mimeographed.
- Reifschneider, D., R. Tetlow, and J. Williams. 1999. "Aggregate Disturbances, Monetary Policy and the Macroeconomy: The FRB/US Perspective." *Federal Reserve Bulletin* (January): 1–19.

- Rotemberg, J. and M. Woodford. 1998. "An Optimization-based Econometric Framework for the Evaluation of Monetary Policy: Expanded Version." Technical working paper t0233. Cambridge, Mass.: National Bureau of Economic Research.
- Sims, C. 1992. "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy." *European Economic Review* 36: 975–1011.
- Sims, C. and T. Zha. 1999. "Error Bands for Impulse Responses." *Econometrica* 67(5): 1113–56.
- Smets, F. and R. Wouters. 2002. "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area." Working paper 171. Frankfurt: European Central Bank.
- . 2003. "Comparing Shocks and Frictions in U.S. and Euro Area Business Cycles: A Bayesian DSGE Approach." Frankfurt: European Central Bank. Mimeographed.
- Sorensen, B.E. and O. Yosha. 1998. "International Risk Sharing and European Monetary Unification." *Journal of International Economics* 45: 211–38.
- Stock, J.H. and M.W. Watson. 2003. "Has the Business Cycle Changed and Why?" In *NBER Macroeconomics Annual 2002*, edited by M. Gertler and K. Rogoff, 159–218. MIT Press.
- Taylor, J.B. 1995. "The Monetary Transmission Mechanism: An Empirical Framework." *Journal of Economic Perspectives* 9(4): 11–26.
- Van Els, P., A. Locarno, J. Morgan, and J.P. Villetelle. 2001. "Monetary Policy Transmission in the Euro Area: What Do Aggregate and National Structural Models Tell Us?" Working paper 94. Frankfurt: European Central Bank.
- . 2003. "The Effects of Monetary Policy in the Euro Area: Evidence from Structural Macroeconomic Models." In *Monetary Policy Transmission in the Euro Area*, edited by I. Angeloni, A. Kashyap, and B. Mojon. Cambridge University Press: 91–106.
- Yun, T. 1996. "Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles." *Journal of Monetary Economics* 37: 345–370.