This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: NBER Macroeconomics Annual 2008, Volume 23

Volume Author/Editor: Daron Acemoglu, Kenneth Rogoff and Michael Woodford, editors

Volume ISBN: 0-226-00204-7

Volume URL: http://www.nber.org/books/acem08-1

Conference Date: April 4-5, 2008

Publication Date: April 2009

Chapter Title: Comment on "How Has the Euro Changed the Monetary Transmission Mechanism?"

Chapter Author: Harald Uhlig

Chapter URL: http://www.nber.org/chapters/c7276

Chapter pages in book: (p. 141 - 152)

Comment

Harald Uhlig, *University of Chicago*

The paper by Boivin, Giannoni, and Mojon seeks to understand the transmission mechanism of monetary policy in the euro area and its constituent countries, document its change since the creation of the euro, and provide a structural interpretation by means of an open-economy model. To do so, it is building on state-of-the-art modeling techniques, most notably Bernanke, Boivin, and Eliasz's (2005) factor-augmented vector autoregressive (FAVAR) approach for the empirical part and Ferrero, Gertler, and Svensson's (forthcoming) open-economy dynamic stochastic general equilibrium (DSGE) model for the structural interpretation. The authors combine both with several innovations, well described in the paper, most notably adding a risk premium on intra-area exchange rates. They report estimated responses to monetary policy largely consistent with conventional wisdom. They document that the creation of the euro has contributed to a widespread reduction in the effect of monetary policy shocks. They interpret this as stemming not only from the adoption of a single currency but also from European Central Bank policy, shifting toward a more aggressive response to inflation and output.

Boivin et al.'s paper exemplifies the best of research that is currently done at central banks as well as in a number of academic departments, seeking to understand aggregate fluctuations and the role of monetary policy from both an empirical and a theoretical perspective. These approaches have started to replace the educated guesses with a serious analysis based on state-of-the-art modeling as the starting point for policy debates. That, in principle, is a good development.

Therefore, I hope that the approach taken here is right. But I fear that severe problems remain and that the route taken here is not yet convincing enough for others to follow. Below I shall explain why, including material found subsequently to my presentation in Boston. Much of

what is stated here may apply with equal force to the predecessors on which the paper at hand is built, and that may seem like a good defense for the authors. But this is their paper in the end, and it is their choice which methodology to apply. Therefore, it is only fair to raise these points here.

I need to warn the reader that this is a discussion. My aim shall be to throw up some challenges and questions and to provoke further thinking on some of these issues. Whether these are fatal flaws or whether all this can be repaired or whether everything is all right after all is something that future research urgently needs to clarify before this approach should be put to wider use. With this disclaimer, let me get in medias res.

I.   The FAVAR Model

There are three basic premises of the empirical approach. First, there is considerable comovement in the selected macroeconomic time series so that their most relevant dynamics is captured by a few factors. Second, the strategy here correctly captures the dynamics associated with monetary policy shocks and correctly identifies their effects. Third, the data are sufficiently informative about the changing impact of monetary policy after the introduction of the euro. I am skeptical about all three.

II.   Is There Comovement in European Data?

The idea that macroeconomic variables comove has considerable appeal in the United States, but perhaps less so in Europe, with its diverse set of countries. Nonetheless, the R^2's reported by Boivin et al. in their table 1 seem impressive and convincing.

But I was still skeptical. If indeed a few factors explain most of what is going on, then the sum of the few largest eigenvalues of the variance-covariance matrix of the data should be near the entire sum of all eigenvalues: that ratio is essentially the R^2's of all variables on the factors corresponding to these eigenvalues. In fact, one would want more: one would want that sum to be considerably larger than in an artificial data set, generated with the same univariate autocovariance structure as in the data, but no comovement among the artificially generated series.

So, I did the following (and I am grateful to the authors for sharing their data set with me to do this). I transformed the data from 1987:Q1 to 2007:Q3 by taking the difference of the log of the current value and

its fourth lag and multiplying by 100, except for interest rates, unemployment rates, and capacity utilization: that way, all data are in percents. This appears to be the transformation chosen by the authors. I call this my baseline data set. I calculated eigenvalues in three ways. First, I took the eigenvalues of the variance-covariance matrix of the baseline data set, summing the largest and calculating the ratio of those partial sums to the total sum. Next, I took the residuals from a regression of the data on current oil and short-term interest rates, that is, series 1 and 243, and a constant and calculated the eigenvalues from the variance-covariance matrix of these residuals (and, as an aside, that seemed to me to be a simpler approach than what the authors have done). Finally, I rescaled all time series to have the same standard deviations before calculating the regression and the eigenvalues of the residuals: from discussions with the authors, it may be that this is closest to the route they have chosen. The results can be seen in figure 1, which lists the number of factors (or largest eigenvalues) on the x axis and the fraction of the total sum of eigenvalues on the y axis. For the x axis I stopped at 30 factors, although there would be 243 (or 83) in principle. One can see that 11 factors in the nonrescaled version explain about
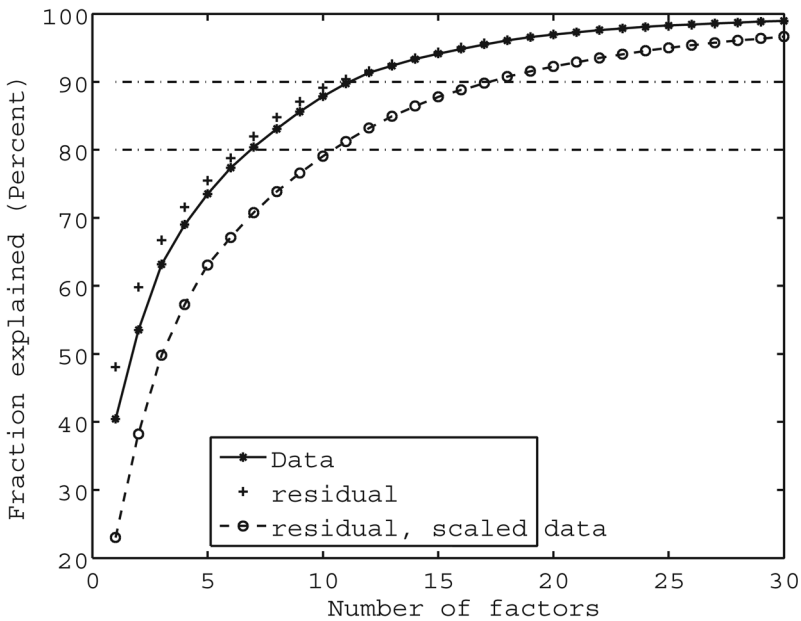

Fig. 1.   Calculated factors and their contribution to overall variance. Three methods of calculating eigenvalues. Authors' original data. This appears to look good.

90%, seven factors get you to about 80% (coinciding roughly with the individual series results in table 1 of the authors), and five factors (think: above and beyond short-term rates and oil) explain about 75%. This initially looks like good news for the approach taken by the authors.

Next, I calculated the first-order autocorrelations of my baseline data set. I then generated an artificial data set as a set of independent AR(1) processes, driven by normally distributed shocks and with the calculated autocorrelations, starting at zero (rather than a draw from the stationary distribution) and rescaled, so that each artificial series has the same standard deviation as the corresponding series in the data. I redid the exact same calculation of the contribution of the factors as above, using the new artificial data series 1 and 243 as regressors: while they have the same autocorrelation as the original data series, there is obviously no reason to expect them to have any explanatory power for the other series. In fact, in the artificial data set, there is no genuine comovement among the series at all.

The result for the artificial data set can be seen in figure 2. I would have expected that figure to be quite different from figure 1 and the
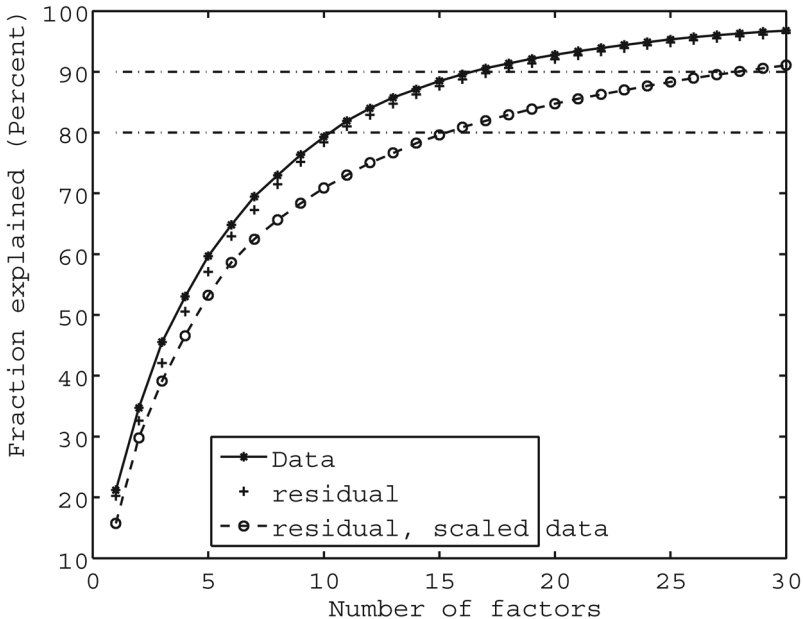

Fig. 2. Like fig. 1 but applied to artificial data: independent AR(1)'s, with autoregressive coefficients distributed as in the original data. This figure is not much different from fig. 1, even though there are no "true" factors in the artificial data. Thus perhaps in the original data, too, the true factors may account for much less comovement in the original data than fig. 1 or the authors' calculations would lead one to believe.

factors with the largest eigenvalues to explain considerably less than in the original data set. But the figures look surprisingly and uncomfortably alike. When I first saw a first version of this figure, I thought that it had to be due to a programming error, accidentally storing the figure coming from the data. But it is really the figure coming from the random data. Yes, there are differences. One factor explains as much now, for either of the three methods. It takes a few more factors to get to the same fraction of variance explained. At five factors for the residual, one is at about 60% rather than 75%. Seven factors deliver about 70% for the baseline random data rather than 80% in the original baseline data. And 12 factors are at 85% rather than 90%. For the residuals from the scaled data, the differences are even somewhat bigger. New random draws will generate slightly different pictures anyhow. "Slightly" is important here. The differences from figure 1, while there, remain strikingly small.

The reason is easy to explain but perhaps tricky to formalize. There is considerable autocorrelation in the data. Figure 3 shows the autocorrelation coefficients, calculated by ordinary least squares and sorted by size: many are close to unity. With persistent roots, deviations from the mean will linger for many periods. Thus, the calculated correlation
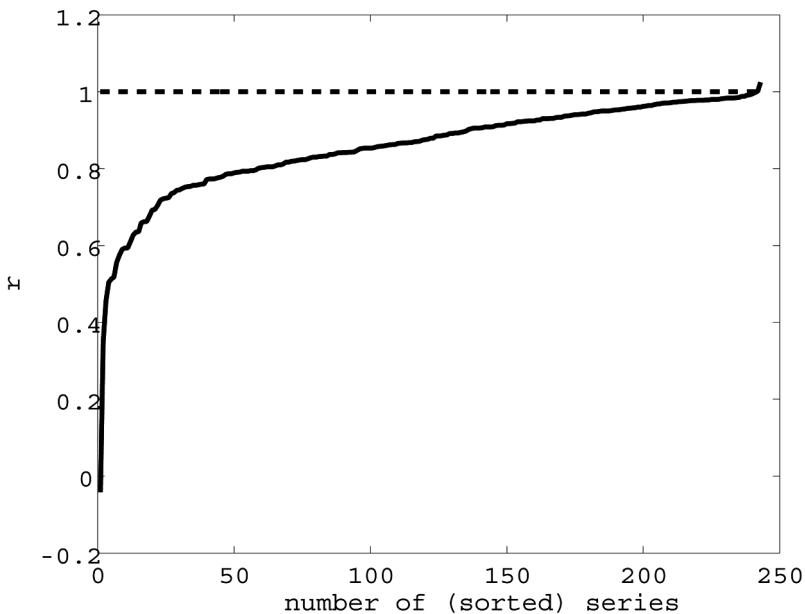

Fig. 3. Distribution of the AR(1) coefficients in the original data, when fitting univariate AR(1)'s to each series. The artificial data for fig. 2 were created as independent AR(1)'s, with the same distribution of AR(1) coefficients.

of two series with persistent roots may easily appear to be large in a finite sample, even though there is none asymptotically. The factors extracted from a finite sample interpret these large correlations as comovements, even though there is none. It all works nicely asymptotically; it just does not work in the short sample at hand and with the large autocorrelations that are in the data. There may be ways around this problem, for example, by prewhitening the series or, at the least, by calculating the factors from the residuals of univariate AR(1) regressions. But this is not what the authors appear to have done.

So in sum, I fear that the approach taken and the evidence presented by the authors are quite consistent with a world in which there is no comovement among the series at all, and they are probably perfectly consistent with a world in which only very few factors matter at the European scale, but explaining considerably less than what the authors make us believe. And without such comovement or too little variation explained by too few factors, the approach has severe problems.

III.   Are Monetary Policy Shocks Identified and Identified Correctly?

But let me give Boivin et al. the benefit of the doubt and hope that my arguments or calculations turn out to be somehow incorrect or not appropriate. That is, suppose that the authors did indeed capture the key comovements and 80% of the variance in the data with their seven factors, including interest rates and oil, even if the sample was truly large. Did they correctly identify monetary policy shocks? I have my doubts.

For starters, it may be that all the movements due to monetary policy shocks have dropped from the sample, once one concentrates on the movement explained by the factors. Cochrane (1994) and many others have argued that monetary policy shocks explain no more than 20% of the movement in the data. It could be that much or even all of that is in the 20% not explained by the leading factors. It is easy to see how this can happen when extracting factors in an unrestricted manner. The authors smartly include the key monetary policy instrument in their factors, but even then, it could happen if the majority of the interest rate movements are not due to monetary policy shocks and if other parts of the movement in interest rates get captured by the seven-factor dynamics and across-variable correlation.

To be more specific, it is worrisome that the fractions explained for M1 and M3 by the factors are among the lowest of all the series (see table 1). We used to think that moving money or moving interest rates

is just as good a tool for a central bank to pick a particular point on the demand curve for money. But table 1 would have to be read as if that demand curve is subject to huge and idiosyncratic fluctuations having nothing to do with the rest of the economy. To put it differently, according to these estimates, money has little or nothing to do with monetary policy and the main movements in aggregate activity, but rather has a life on its own. If you believe this, you have an interesting research agenda at hand.

But even leaving these arguments aside, I seriously wonder whether the approach to identify monetary policy shocks is reasonable. Section IV. A states that it is assumed that "the latent factors … and the oil price inflation … cannot respond contemporaneously to a surprise interest rate change." The argument for this approach is in Bernanke et al. (2005), in which the authors argue that the movement in the factors is movement due to "slow-moving" variables, since any additional systemic movement in the "fast-moving" variables is one-dimensional; they interpret this as being largely explained by the surprise in interest rate movements. But there are no such things as slow-moving variables. After all, all variables have a nonzero one-step-ahead prediction error: they thus move fast with respect to something. The identifying assumption here really is that whatever it is they are reacting to contemporaneously and quickly, it cannot be monetary policy. Why should that be the case? If inflation and employment can suddenly jump a bit because of shifts in market demands, why can they not do so when monetary policy surprisingly changes interest rates?

The defense seems to be that the impulse responses look conventional. But they don't. As figure 1c in Boivin et al.'s paper shows, consumer price index inflation in Germany, France, Italy, and the euro area as a whole tends to move up rather than down after a monetary tightening, and wage inflation moves up in Germany, Italy, and Spain. Additionally, these responses are estimated with a fairly wide error band. The reaction seems to be somewhere between −0.3% and 0.3% in the year following the shock. By contrast, the reaction of GDP is fairly sharp and always down, ranging from −1% or below to about −0.2% in the year following the shock. That seems large compared to the (non)movement in inflation.

A more convincing approach to identification is to employ the conventional wisdom and therefore sign restrictions for identification, as I have proposed in Uhlig (2005). With a panel of macroeconomic time series and a factor approach, as in the paper at hand, there are considerably more sign restrictions that can aid in identification, and the

methodology then provides for considerably sharper bounds as well as reasonable results (see Ahmadi and Uhlig 2008).

IV. Are the Data Informative about the Change after the Introduction of the Euro?

I do not need to answer that question. Boivin et al. themselves provide ample warning in their paper that this is not so. Note in particular that no error bands have been provided to the post-euro responses in figures 1*a*–*c* or the comparison pictures. Be wary of econometricians who draw conclusions by comparing means without telling you the degree of uncertainty! It is a fair guess that it is large. There simply was not much time-series variation in monetary policy since the introduction of the euro. Figure 4 shows what is going on: large and heterogeneous movements in interest rates before the introduction of the euro. Hardly any movements afterward.

The authors are probably happy that the impulse responses did not change too dramatically for several key variables. Unfortunately, there are some in which the responses did change, leading us even further
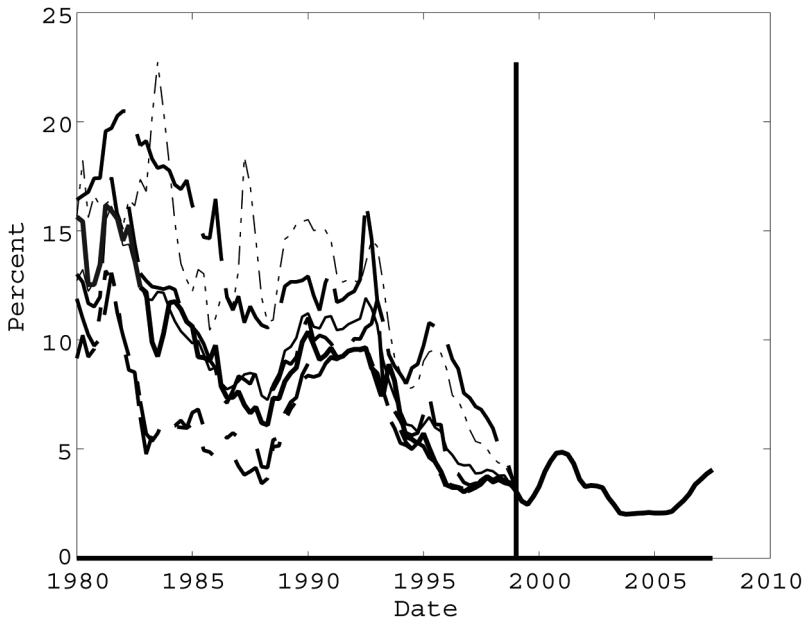

Fig. 4. Short-term interest rates in the EMU, authors' data: euro area, Germany, France, Italy, Spain, Netherlands, and Belgium.

away from conventional wisdom. Consumption moves up after a monetary tightening. M3 moves up substantially now after a monetary tightening, quite in contrast to what happened before the euro.

One explanation within the philosophy of the authors is that post-euro monetary policy shocks identified here are really capturing movements in the stock market. For suppose that there are practically no monetary policy shocks and that monetary policy is instead also reacting to movements in some other fast-moving variable, such as the stock market. Suppose an econometrician knew that and wanted to identify stock market surprise movements above and beyond those of slow-moving variables. That econometrician would have proceeded exactly as the authors did, except that the impulse responses now would have to be interpreted as responses to stock market shocks rather than monetary policy shocks. How can one tell them apart? Again, sign restrictions might help.

V.   The Structural Model

The paper complements the empirical analysis with a structural model that allows one to interpret the data from that vantage point. The key difficulty for this model is to explain the interest rate convergence in figure 4, happening without correspondingly large inflation differences. The authors readily admit this problem in Section V.C.1, when they write that "the basic version of the model cannot replicate the transmission of monetary policy observed in low-credibility regimes since long-term rates are tightly tied to expected future riskless short-term rates." One possibility would be to scrap the model at this point.

The authors instead invent a clever deus ex machina: shocks to the uncovered interest rate parity (UIP) condition, which furthermore are tied with a key parameter to foreign (or "German") monetary policy shocks (see Sec. V.B.3). Let me put it differently. Most of the interesting action in monetary policy in Europe over the last 20 years is the convergence process seen in figure 4. The authors sweep all that away by an add-on to the UIP condition, which, however, has no further implications for aggregate dynamics. Next, they then seek to study how the changes in monetary policy from the pre-euro regime to the post-euro regime have affected macroeconomic variables. Shouldn't one worry a bit that the baby has already been thrown out with the bath water? There is something really interesting happening here: it is the major big thing in the transition to the euro. We cannot quite put it into the theory, so thus let us ignore it? Shove it into a random shock, leaving everything else unchanged?

I can see the desperation of the authors here, and I laud them for their frankness. Figure 4 is hard to explain within this theory. It is my guess too that it has a lot to do with perceptions of risk and updating the probabilities of membership in the European Monetary Union (EMU). So, having gotten so far in setting up this beautiful model and all, I understand that the quick fix of declaring it to be completely uninteresting and tangential was a way to proceed with the rest. But here is a memo to subsequent research: forget about the rest and instead put this at center stage, to understand the role of changing monetary policy in Europe!

The authors instead plug in reaction coefficients of monetary policy, which are not obtained from the previous empirical exercise and not obtained from estimating the structural model, but instead from another empirical exercise described in Section V.B.2. One has to wonder whether this is consistent with the initial FAVAR approach or with the structural model at hand. Anyway, given that they use different coefficients before and after EMU, they find different quantitative results of their model. This is what the main comparison of pre-euro and post-euro in the paper rests on. Perhaps a more serious subsample stability test, using the structural DSGE model for estimation rather than an auxiliary model (or, at least, providing a link with the tools of indirect inference), would be more convincing.

The connection to the empirical approach of the first half of the paper is rather tenuous, and more tenuous than may appear at first. For example, the identification assumption that all other variables do not react to monetary policy shocks within the period is maintained. But that follows essentially just from an artifact of notation. What is the difference between a monetary policy shock that is happening at the end of a period and a monetary policy shock happening at the beginning of the next? It just depends on the artificial way the continuous time line is broken up into discrete periods. In figure 3, one could equivalently read the impulse responses by moving them to the left by one period and having all variables react within the period to the (beginning-of-period) monetary policy shock. Also, it would be interesting to compare the monetary policy shocks as identified by the DSGE model to the monetary policy shocks as identified by the VAR.

Is this a good model to study the impact of monetary policy in Europe? Note that there is no financial intermediation here. There is no financial sector that could potentially get in trouble from mortgage-backed securities. There is no worry about companies investing less when interest rates rise, because there is no capital and no investment in the model to begin with. High unemployment, the sclerosis of European labor

markets, and the many frictions introduced by fiscal policy and the welfare state are also essentially absent.

The only friction that monetary policy worries about is due to the stickiness of intermediate-goods prices. Fortunately for the model, importers and exporters here are no different from domestic firms. But aren't exchange rates fluctuating substantially more than within-Europe post-euro inflation rates? If sticky prices are the main concern, shouldn't monetary policy perhaps be much more focused on the fact that the dollar-euro exchange rate rose from 0.90 to 1.60—and the ensuing distortions in the relationships of fixed dollar-euro prices—rather than the comparatively tiny distortions created by some domestic firms not being able to adjust prices but others can? Put differently, is the main concern of Mercedes that it cannot adjust its prices for cars in Germany whereas Volkswagen can? Or is it more important to them that their margins erode as they are trying to compete at German wage levels against U.S.-based car companies, which pay U.S. wages while the value of the dollar collapses? And if that is the more important issue, how much might that have played a role in the transition to the EMU for monetary policy, as a common currency got established, and for monetary policy overall?

In sum, the model here is at the current edge of quantitative research on monetary policy. But I am afraid that several of the most interesting features, which really matter for monetary policy and really matter for the transition to the EMU, have been tossed out before the analysis has even begun. And if so, then the problems with this approach are severe indeed. We then ought to pursue different models to address the main issues of monetary policy.

VI.   Conclusions

I enjoyed discussing Boivin et al.'s paper. I really did: some may remember. The paper is a serious, honest, hard-core analysis at the current edge of research and among the best that one can find on the issue at hand, building on the best tools available. I applaud the authors for what they have accomplished: this is no small feat. I am hoping that the approach is correct and that we can build serious monetary policy discussions on this basis.

But I fear that severe problems remain and that the route taken here is not yet convincing enough for others to follow. I have outlined why.

I fear that the approach taken and the evidence presented by the authors are quite consistent with a world in which there is no comovement among the series at all; they are probably perfectly consistent with

a world in which only very few factors matter at the European scale, but explaining considerably less than what the authors make us believe. And without such comovement or too little variation explained by too few factors, the approach has severe problems. Even if there are common factors, the monetary policy shocks may be incorrectly identified; witness the price puzzles and the sharp reaction of output compared to the muted reaction of inflation.

The theory needs to introduce the deus ex machina of exogenous shocks to UIP in order to explain what must be the key feature of monetary policy in Europe, namely, convergence of interest rates (see fig. 4). Several other key issues of central importance to monetary policy play no role in the theory either. And if so, then the problems with this approach are severe indeed. We then ought to pursue different models to address the main issues of monetary policy.

Finally, there is a disconnect between the theory and the empirics.

All this is cause for concern. Whether these are fatal flaws or whether all this can be repaired or whether everything is all right after all is something that future research urgently needs to clarify before this approach should be put to wider use.

References

Ahmadi, Pooyan Amir, and Harald Uhlig. 2008. "Measuring the Dynamic Effects of Monetary Policy Shocks: A Bayesian FAVAR Approach with Sign Restrictions." Manuscript, Humboldt University Berlin and University of Chicago.

Bernanke, Ben, Jean Boivin, and Piotr Eliasz. 2005. "Measuring Monetary Policy: A Factor Augmented Vector Autoregressive (FAVAR) Approach." *Quarterly Journal of Economics* 120, no. 1:387–422.

Cochrane, John. 1994. "Shocks." *Carnegie-Rochester Conference Series on Public Policy* 41:295–364.

Ferrero, Andrea, Mark Gertler, and Lars E. O. Svensson. Forthcoming. "Current Account Dynamics and Monetary Policy." In *International Dimensions of Monetary Policy*, ed. Jordi Galí and Mark Gertler. Chicago: University of Chicago Press.

Uhlig, Harald. 2005. "What Are the Effects of a Shock to Monetary Policy? Results from an Agnostic Identification Procedure." *Journal of Monetary Economics* 52:381–419.