

Stronger Evidence of Long-Run Neutrality: A Comment on Bernanke and Mihov*

Paolo Giordani[†]

Abstract

Few propositions in macroeconomics are less controversial than long-run money neutrality, yet clear and robust empirical support has not been found in time series studies. Bernanke and Mihov (1998) are comparatively successful in this hunt, but their output response to monetary policy shocks remains stubbornly persistent. This paper argues that the omission of a measure of output gap from the VAR estimated by Bernanke and Mihov lies at the heart of this "excessive" persistence. In the theoretical framework of a New Keynesian model similar to that of Svensson (1997) and Clarida, Gali and Gertler (1999), I prove that this omission induces persistence overestimation under relatively mild assumptions. The inclusion of a proxy for the output gap in the VAR is then shown to drastically increase the evidence for long-run money neutrality on US data, as predicted by the theoretical analysis.

Keywords: long-run money neutrality, technology shocks, output gap, VAR misspecification.

JEL: E 31, E 52, E 58.

*I thank Ben Bernanke, Lars Svensson and Paul Söderlind for comments. Financial help from the Wallanders and Hedelius Foundation is gratefully acknowledged.

[†]Stockholm School of Economics, Department of Economics. Box 6501, SE-11383, Stockholm, Sweden. Phone: +46 8 7369252. Fax: +46 8 313207. E-mail: nepgi@hhs.se.

1 Introduction

Few propositions in macroeconomics are less controversial than long-run money neutrality (henceforth LRN), the assertion that a monetary contraction (expansion) has no effect on employment and real output in the long run. Yet, clear and robust empirical evidence in favor of LRN seems hard to gather in time series studies. Bernanke and Mihov (1998b) address this problem and go a long way in clarifying the errors that had caused previous empirical research to reject LRN. Bernanke and Mihov (henceforth BM) start with a strong prior in favor of LRN: they are out to find support for it. But although their work is more successful than previous research in this hunt, their success lies in not providing a clear rejection of LRN rather than in providing strong evidence in its favor: a null hypothesis of large long-run effects would not be rejected either. Their strategy consists in isolating a monetary policy shock (henceforth *MP* shock) by way of a structural VAR, and in computing the response of output to that shock. LRN implies that the response should converge to zero as the time horizon grows. There is a number of interesting methodological advances in BM; among these, they convincingly argue that choosing a broad monetary aggregate (monetary base or broader) as the instrument of monetary policy is a poor choice which leads to huge mistakes on the issue of LRN. But in spite of these improvements, the response of output is not back at zero after ten years. In their own words:

”We find that the liquidity effect and LRN propositions are not inconsistent with our approach to identifying monetary policy, although the point estimates for the output response imply what might be viewed as an ”excessive” degree of persistence from the perspective of the LRN hypothesis”.¹

”At four years the output response is no longer statistically different from zero. It is a bit troubling however, that the point estimates of the output response remain above zero even at the ten-year horizon”.²

If LRN is so hard to nail down, should we take a pause and doubt it? After all, ”hysteresis” is a word which many economists are not ashamed of using in other contexts. Mankiw (2000) does raise the question and does use the word in a section titled ”Hysteresis?”. In fact, Mankiw reads BM’s results as a ringing bell in favor of a more agnostic position on LRN, rather than as evidence for it. In his words:

”One question about which I remain agnostic is whether the natural rate hypothesis is true or whether some form of hysteresis causes monetary shocks to have long-lasting effects on unemployment. [...] Bernanke and Mihov estimate a structural vector autoregression and

¹Pg. 154.

²Pg. 171.

present the impulse response functions for real GDP in response to a monetary policy shock. (See their Figure III) Their estimated impulse response function does not die out toward zero, as is required by long-run neutrality. Instead, the point estimates imply a large impact of monetary policy on GDP even after ten years. Bernanke and Mihov don't emphasize this fact because the standard errors rise with the time horizon. Thus, if we look out far enough, the estimated impact becomes statistically insignificant. But if one does not approach the data with a prior view favoring long-run neutrality, one would not leave the data with that posterior. The data's best guess is that monetary shocks leave permanent scars on the economy".

In what follows I propose an explanation for the persistent effect of *MP* shocks which assumes LRN. The question I ask is: if LRN holds, why does BM's VAR produce such a persistent response of output to *MP* shocks? My answer is fully based on economic theory, indeed in a model that incorporates LRN and a short interest rate as the main operating tool of modern central banking. The model, which is borrowed from Svensson (1997) with minor additions, is designed to capture the essential features of the monetary transmission mechanisms, as outlined in Svensson (1997 and 2000a) and in Clarida, Gali and Gertler (1999). In the model the output gap plays an important role: it affects inflation and it belongs to the reduced form reaction function of the monetary authority (which turns out to be a simple Taylor rule). This paper shows that, under reasonable parameter values, the omission of a measure of output gap from the model results in overestimating the persistence of the output response to monetary policy shocks. Since BM do not include a measure of output gap in their VAR, this analysis is potentially a valid explanation for their results. My next step is then to move to the data, to check whether adding a measure of output gap to the VAR does produce stronger evidence in favor of LRN. The result I propose is that the evidence in favor of LRN becomes much stronger once a measure of output gap is added to the VAR, just as predicted by the model. I am eager to stress at this early point that this large improvement (from the prospect of BM) relies entirely on the introduction of one additional variable (a measure of the output gap), as suggested by theory: everything else, including identification, is the same for both specifications.

The intuition for why the omission of an output gap measure results in an overly persistent response of output to *MP* shocks is as follows. The output gap belongs to the reduced form monetary policy rule, in such a way that, for a given unexpected increase in output, the monetary authority raises the interest rate by more if the increase is due to an aggregate demand shock than if it is due to a technology shock, defined as a shock that changes potential output (this result is rather standard in the literature); since the output gap is omitted, when output increases, the VAR sometimes observes a large increase in the interest rate (*AD* shock) and sometimes a smaller one (technology shock), and therefore estimates that the interest rate response to an unforecasted output increase is somewhere in between. It follows that when a positive technology

shock increases output, the VAR expects the interest rate to raise by more than it actually does, and interprets the difference as an expansionary *MP* shock. Therefore the *MP* shocks estimated by the econometrician are contaminated by technology shocks. If technology shocks have a longer lasting effect on output than *MP* shocks, as commonly believed, the response of output to a *MP* shock in the misspecified system looks more persistent than it actually is.

The rest of the paper proceeds as follows. Section 2 develops this intuition formally, by presenting the model and showing the consequences of omitting a measure of output gap in the VAR, both analytically and through simulations. Section 3 tests the implications of the analysis carried out in Section 2. The VAR used by BM is compared with one that, *ceteris paribus*, adds a measure of output gap to the non-policy variables, finding that the evidence in favor of LRN is dramatically increased, as predicted. Section 4 concludes.

2 Model and consequences of VAR misspecification

In order to understand the consequences of omitting a measure of output gap from the VAR, I refer to a simple model which replicates, with some additions, the one in Svensson (1997). Reference to this model will allow for a clear and rigorous exposition. However, the main result I am after (overestimated degree of persistence) is grounded on an intuition which holds for a more general class of models.

Svensson (1997) presents a model designed to capture some key features of the transmission mechanism of monetary policy. In fact, it is more generally a model of business cycle fluctuations. The same model is used in Rudebusch and Svensson (1999), in Judd and Rudebusch (1998) and, extended to a small open economy, in Ball (1999). A forward-looking version appears in Clarida, Gali and Gertler (1999) and in Svensson (2000a and 2000b). Romer (2000) presents the same model as an improvement over the traditional *IS-LM*. The model consists of an *IS* equation, a Phillips curve and a Taylor rule obtained from the monetary authority's optimization problem. This core three-equation structure is shared by many recent New-Keynesians models for monetary policy analysis. The model assumes that the monetary authority uses a short interest rate as its main operating target.

The *IS* relation is given by

$$y_{t+1}^g = \beta_y y_t^g - \beta_r (i_t - \pi_t) + \epsilon_{t+1}^{AD}, \quad (1)$$

where i_t is a short-term interest rate set by the monetary authority, π is inflation and y^g is the output gap, defined as $y_t^g = Y_t - Y_t^N$, where Y_t is the log of output and Y_t^N the log of potential (or "natural") output. Potential output is assumed

to follow an exogenous AR(1) process³

$$Y_{t+1}^N = \rho Y_t^N + \epsilon_{t+1}^N. \quad (2)$$

The Phillips curve is modelled as

$$\pi_{t+1} = \pi_t + \alpha_y y_t^g + \epsilon_{t+1}^{CP}. \quad (3)$$

All shocks are *iid*.⁴ They are labelled: aggregate demand shock, technology shock and cost-push shock. Denote their standard deviations by σ_{AD} , σ_N , σ_{CP} . The model is supplemented by a loss function for the monetary authority of the standard type

$$L_t = E_t \sum_{i=0}^{\infty} \beta^i [\lambda (y_{t+i}^g)^2 + (\pi_{t+i} - \pi^*)^2]. \quad (4)$$

The solution takes the form of a Taylor rule (See Svensson (1997) for the closed-form solution). Since the model is backward-looking the discretionary solution and the commitment solution are the same:

$$i_t = \gamma_\pi \pi_t + \gamma_y y_t^g. \quad (5)$$

A monetary policy shock can be added by supposing that the Taylor rule is not followed deterministically. In that case, the shock ϵ^{MP} with std σ_{MP} is added to the Taylor rule. *MP* shocks affect neither output nor inflation within the period.

Money does not have an interesting role to play in this model: it does not appear in the *IS* and Phillips equations, and i being the policy instrument, money supply is perfectly elastic. Therefore, for a given money demand equation, the quantity of money can be determined residually, after solving the rest of the model, and the omission of money does not result in any misspecification.⁵

In this model, *MP* shocks can be identified by the assumption that they do not affect output, potential output and inflation within the period. Thus *MP* shocks can be recovered with a triangular identification, where all variables except money are ordered before the interest rate (within these limits, the ordering of the other variables is irrelevant for the identification of *MP* shocks). The identifying assumption that *MP* shocks do not affect output and inflation within the period is a rather uncontroversial one, and is used extensively in the literature (see, e.g., Christiano et al., 1998, and Rudebusch and Svensson, 1998). Therefore, in this respect, BM's assumptions are fully in line with the model. On the other hand, the model does not justify the omission of the output gap (or potential output) from the VAR, as in BM and in much of the literature.

³Svensson (1997) makes no assumption about potential output. I follow Svensson (2000a and 2000b) in assuming an AR(1) process.

⁴The assumption of *iid* shocks is not particularly restrictive, as more lags can be added to equations (1) to (3) without any difficulty.

⁵Given these premises, it is not surprising that the literature adopting models of this type does not specify a money demand equation. However, Leeper (2000) expresses some doubts on the appropriateness of omitting money from the analysis.

2.1 Omitting a variable from the VAR and finding excessive persistence

Let the DGP be given by equations (1) – (3) and (5). Suppose that we estimate a VAR including: output, inflation, interest rate and money, with MP shocks identified as interest rate shocks, and interest rate ordered second to last (money last). The identification would be in accordance with the model if output gap or potential output were added to the VAR, ordered before the interest rate. The message of this paper is that omitting a measure of output gap from the VAR results in excessive (overestimated) persistence.

A perspective on excessive persistence can be gained by inspecting the retrieved MP shocks. In fact, though the Taylor rule is assumed deterministic, the econometrician will retrieve a positive std of the structural MP shocks. This point is proved formally in Giordani (2001).⁶ However, the reason for the overestimation of the variance of MP shocks is intuitively clear: since the output gap is omitted from the estimated Taylor rule, the fit deteriorates (since the output gap is not perfectly collinear with inflation, output and lags of all variables). The intuition is that since the interest rate does not react in the same way to technology and AD shocks, when a movement in output (of a given amount) is observed, the misspecified VAR sometimes registers a large change in the interest rate (when the movement is caused by an AD shock) and sometimes a smaller (possibly null) change (when caused by a technology shock) and is tricked into interpreting this as random behavior of the monetary authority. It also estimates that the interest rate response to an unforecasted output increase is somewhere in between.⁷ It follows that when a positive technology shock pushes output up, the VAR expects the interest rate to raise by more than it actually does, and interprets the difference as an expansionary MP shock. Therefore the MP shocks estimated by the econometrician are contaminated by technology shocks; more precisely, they are negatively correlated with technology shocks. If technology shocks have a longer-lasting effect on output than MP shocks, as commonly believed, the response of output to a MP shock in the misspecified system will be more persistent than in reality. Giordani (2001) shows that the covariance is given by

$$\text{cov}(\epsilon_{t+1}^{*MP}, \epsilon_{t+1}^N) = -\gamma_y \frac{\sigma_{AD}^2 \sigma_N^2}{\sigma_{AD}^2 + \sigma_N^2} < 0. \quad (6)$$

2.2 Simulation

So far I have presented analytical arguments predicting that the omission of a measure of output gap will lead to excessive persistence of the output response to the estimated MP shocks. To develop a quantitative measure of the importance of this effect, I present a simulation. The model parameters are set

⁶The std of the retrieved MP shocks is $\sigma_{MP}^{*2} = \gamma_y^2 \frac{\sigma_{AD}^2 \sigma_N^2}{\sigma_{AD}^2 + \sigma_N^2} > 0$.

⁷In this model the interest rate does not react at all to technology shocks, but the intuition holds as long as the output gap belongs in the reduced form policy function.

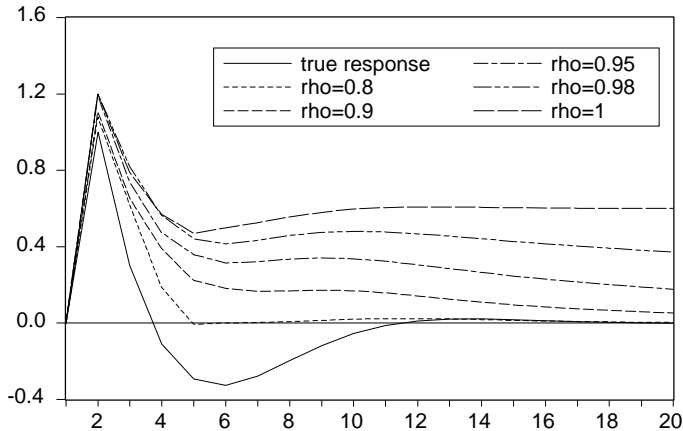


Figure 1: Impulse response of output to an expansionary MP shock in the correctly specified model and in the misspecified model for various degree of persistence of potential output.

as in Ball (1999): $\alpha_y = 0.4$, $\beta_y = 0.8$, $\beta_r = 1$. The standard deviations are $\sigma_{AD} = \sigma_{CP} = \sigma_N = \sigma_{MP} = 1$. Results generalize without surprises to different parameters. The parameters in the Taylor function are set to $\gamma_y = 0.5$, $\gamma_\pi = 1.5$.⁸ When $\sigma_{MP} > 0$, the misspecified model is more easily estimated numerically and allowing for several lags (here 4), for reasons explained in Giordani (2001). I abstract from parameter uncertainty by setting the number of observations to 50000. Figure 1 shows the true impulse response of output to a MP shock together with the impulse responses retrieved by the misspecified VAR, for increasing values of the autoregressive parameter ρ . As implied by the theoretical analysis, the degree of excessive persistence increases with ρ . Most researchers' prior is that ρ is either one or close to one, which in this model could explain why the response is so slow in returning to zero.⁹

3 Moderate persistence in US data

The natural next step is to verify the accuracy of these predictions on US data. To iterate, the prediction is that the support for LRN will be stronger if a measure of output gap is included in the VAR, among those that BM call "non-policy

⁸These parameters are commonly adopted. However, the optimal values of the parameters of the Taylor function has a closed form solution given in Svensson (1997).

⁹A Monte Carlo exercise supports Gali (1998)'s argument that, with a small sample (200 observations in the experiment) and in the presence of (close to) unit root output, estimating the VAR in levels will lead to underestimating the persistence of the response. However, this effect is small.

variables". Notice that the model presented above supports the identification assumption, adopted by BM, that monetary policy does not affect non-policy variables within the period, with one exception and one clarification. The exception is that the broad monetary aggregate $M2$ should be ordered after the policy variables, not before. The clarification is on the role of the commodity price index, which does not appear in the model. Giordani (2001) argues that the commodity price index¹⁰ is useful in solving the price puzzle because it is correlated with popular measures of the output gap (in the period considered by BM (1966-1996), the correlation is 0.43 with capacity utilization¹¹ and -0.23 with unemployment), and that it is no longer essential if a better proxy of the output gap is included in the VAR. The results that follow support this claim: while BM's results (in terms of persistence of the output response) worsen considerably if commodity price (henceforth $Pcom$) is omitted, this omission has no consequences if the output gap is included.

I will simplify exposition by reducing the policy variables to one only: the federal funds rate. Bernanke and Mihov (1998a and 1998b), who also consider total reserves and nonborrowed reserves, in fact conclude that this is a good approximation for most of the post-1965 period.

I will reproduce the first two graphs of Figure III in BM (1998b), namely the response of real GDP to a MP shock and the p-value for the null that the response is zero after a given number of periods; but I will do this comparing the results of two VARs:

1. A VAR identical (in variables and identification) to the one used by BM (except for the simplification of the policy block to the federal funds rate): real GDP, inflation, $Pcom$, $M2$, federal funds rate.¹²
2. The VAR suggested by the model, which is the same as above (same identification ordering) but adding capacity utilization (among the non-policy variables) and without $M2$ and $Pcom$. The results obtained by this second VAR are very similar to one in which $M2$ and $Pcom$ are included. I present the smaller VAR because of Occam's razor principle (here: if a small model performs as well as a large one, use the small) and to avoid Sims' objection to ordering $Pcom$ before the policy instrument.

Data are quarterly, sample 1966Q1-1998Q4.¹³ The VARs have three lags.¹⁴ Figure 2 graphs the impulse response of GDP to an expansionary MP shock

¹⁰The commodity price index I choose is the series used in Christiano et al. (1998), which is a leading indicator published by the Department of Commerce, arguably the most useful in correcting the price puzzle. This series does not have an upward trend, as the word "price" may suggest and as other similarly denominated series do (such as the international commodity price index used by Favero, 2000).

¹¹Capacity utilization for the industrial sector is published by the Federal Reserve Board, also available at FRED (www.stls.frb.org/fred/data/business/comfg).

¹² $M2$, in BM (1998b) and here, is deflated by the DGP deflator and logged (the same is true of GDP). I use inflation, rather than prices, to improve efficiency in estimation.

¹³BM use monthly data (1966-1996), for which capacity utilization is not available.

¹⁴The HQ criterion suggests two lags for the first VAR and three for the second. BM (1998b) use seven lags on monthly data.

in the two models (for 40 quarters, as in BM), together with 68% error bands (approximately two standard deviations) for the second model.¹⁵ Figure 3 plots the p-values for the two hypothesis that the response is zero after a given number of periods in the first VAR (first hypothesis) and in the second VAR. The results

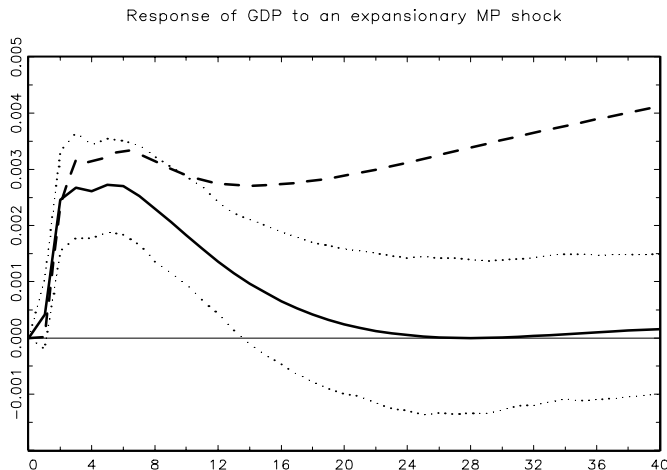


Figure 2: Impulse response of GDP to an expansionary *MP* shock in the second VAR (capacity, GDP, inflation, federal funds rate)—continuous line, error bands in dotted lines—and in the first VAR (GDP, inflation, Pcom, M2, federal funds rate)—dashed line.

from the first VAR are slightly different from those obtained by BM (see their Figure III): the p-value raises above 0.1 sooner, but at longer horizons provide less support for LRN. On the other hand, the second VAR drastically increases the support for LRN, while preserving very significant short-run effects. Notice the similarities between Figure 1 and Figure 2: in the VAR that includes capacity utilization, the response is smaller, but peaks at approximately the same time, and is shorter-lived. This baseline result, namely that simply including a measure of output gap in the VAR drastically reduces the persistence of the output response, is robust to a wide range of modifications including: changing lag length, adding Pcom and M2 to the second VAR, differencing GDP in estimation (as suggested by Gali, 1998), using CPI rather than GDP deflator, shortening the sample, using unemployment instead of capacity utilization as a

¹⁵Error bands and p-values are obtained by Monte Carlo numerical integration (10000 draws) with the Bayesian procedure outlined in Doan (1992). Error bands are obtained leaving out the lowest and highest 16% at each step.

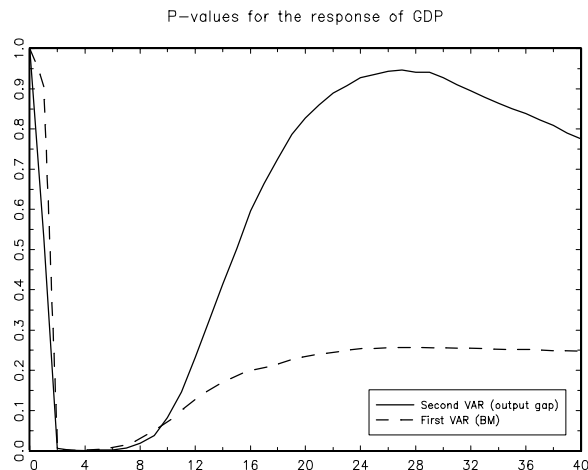


Figure 3: P-values for the null that the response of output at a given horizon is zero for the first VAR (BM) and the second VAR (capacity, output, inflation, ff).

measure of output gap. The exclusion of Pcom, irrelevant if capacity is included, leads to much weaker evidence in support of LRN in the first VAR (the p-value at lag 40 is 0.07).

The one presented here is not the first VAR that does not display "excessive" persistence. The impulse response in, for example, Rotemberg and Woodford (1998)—who detrend output—and Rudebush and Svensson (1999)—who include the CBO measure of output gap—also do not. The contribution of this paper is to rationalize their findings, explaining what drives the differences in results across apparently similar VARs.

4 Conclusions

Few propositions in macroeconomics are less controversial than long-run money neutrality, yet clear and robust empirical support has not been found in time series studies. Have economists been too stubborn in interpreting evidence against LRN as a failure of the empirical methodology rather than as evidence against the hypothesis? This paper suggests that they have not. Commenting on work by Bernanke and Mihov (1998b), and sharing their main assumptions on monetary policy identification, the paper argues that the omission of a measure of output gap (or potential output) from the VAR estimated by Bernanke and

Mihov lies at the heart of the "excessive" persistence of the output response to *MP* shocks. In a theoretical framework, this omission is proven to induce persistence overestimation under relatively mild assumptions. The inclusion of a proxy for the output gap in the VAR is then shown to drastically increase the evidence for long-run neutrality on US data, as predicted by the theoretical analysis.

References

- [1] Ball, L., 1999, Policy rules for open economies, in: John B. Taylor, Monetary Policy Rules, (University of Chicago Press, Chicago).
- [2] Bernanke, B. S., I. Mihov, 1998a, Measuring monetary policy, Quarterly Journal of Economics, 113, 869-902.
- [3] Bernanke, B. S., I. Mihov, 1998b, The liquidity effect and long-run neutrality, Carnegie-Rochester Conference Series on Public Policy 49, 149-194.
- [4] Christiano, L.J., M. Eichenbaum and C. L. Evans, 1998, Monetary policy shocks: What have we learned and to what end?, NBER WP #6400.
- [5] Clarida, R., J. Gali and M. Gertler, 1999, The Science of monetary policy, Journal of Economic Literature 37, 1661-1707.
- [6] Doan, T. A., 1992, RATS User's manual, Estima.
- [7] Favero, C., 2000, Applied macroeconometrics, (Oxford University Press, Oxford).
- [8] Gali, J., 1998, The liquidity effect and long-run neutrality: A comment, Carnegie-Rochester Conference Series on Public Policy 49, 195-206.
- [9] Giordani, P., 2001, An alternative explanation of the price puzzle, Working paper, Stockholm School of Economics. Available at <http://swopec.hhs.se/hastef/abs/hastef0414.htm>.
- [10] Judd, J. P. and G. D. Rudebush, 1998, Taylor's Rule and the Fed:1970-1997, Economic Review, 3, Federal Reserve Bank of San Francisco.
- [11] Leeper, E., 2000, Discussion of Mike Woodford's "A New-Wicksellian framework for the analysis of monetary policy", available at <http://php.indiana.edu/~eleeper/#Papers>.
- [12] Leichter, J. and C. Walsh, 1999, Different economies, common policies: Policy trade-offs under the ECB, working paper. Available at <http://econ.ucsc.edu/~walshc>.
- [13] Mankiw, N. G., 2000, The inexorable and mysterious tradeoff between inflation and unemployment. Available at <http://post.economics.harvard.edu/faculty/mankiw/papers.html>.

- [14] Romer, D., 2000, Keynesian macroeconomics without the LM curve, *Journal of Economic Perspectives*, 14, 149-169.
- [15] Rotemberg, J. J. and M. Woodford, 1998, An optimization-based econometric framework for the evaluation of monetary policy: expanded version, NBER Technical Working Paper 233.
- [16] Rudebusch, G. D. and L. E. O. Svensson, 1999, Policy rules for inflation targeting, in: John B. Taylor *Monetary Policy Rules*, (University of Chicago Press).
- [17] Sims, C., 1992, Interpreting the macroeconomic time series facts: The effects of monetary policy, *European Economic Review*, 36, 2-16.
- [18] Svensson, L. E.O., 1997, Inflation forecast targeting: Implementing and monitoring inflation targets, *European Economic Review*, 41, 1111-1146.
- [19] Svensson, L. E.O., 1999, Inflation targeting as a monetary policy rule, *Journal of Monetary Economics*, 43, 607-654.
- [20] Svensson, L. E.O., 2000a, "Open-economy inflation targeting," *Journal of International Economics*, 50, 155-183.
- [21] Svensson, L. E.O., 2000b, "The zero bound in an open economy: A foolproof way of escaping from a liquidity trap", NBER working paper n.7957.