Working Paper

INTERNATIONAL MONETARY FUND
Does Money Matter for U.S. Inflation?
Evidence from Bayesian VARs

Helge Berger and Pär Österholm
IMF Working Paper

Western Hemisphere Department

Does Money Matter for U.S. Inflation? Evidence from Bayesian VARs∗

Prepared by Helge Berger and Pär Österholm

Authorized for distribution by Robert Rennhack

March 2008

Abstract

We use Bayesian estimation techniques to investigate whether money growth Granger-causes inflation in the United States. We test for Granger-causality out-of-sample and find, perhaps surprisingly given recent theoretical arguments, that including money growth in simple VAR models of inflation does systematically improve out-of-sample forecasting accuracy. This holds for a long forecasting sample 1960-2005, as well for more recent subperiods, including the Volcker and Greenspan eras. However, the contribution of money to inflation forecasting accuracy is quantitatively limited and tends to be smaller in recent subperiods, in particular in models that also include information on real GDP growth and interest rates.

JEL Classification Numbers: E47, E52, E58

Keywords: Out-of-Sample Forecasting, Granger Causality, Monetary Aggregates, Monetary Policy, Volcker, Greenspan

Authors’ E-Mail Addresses: hberger@imf.org; posterholm@imf.org

∗ Both authors thank the IMF for its hospitality. Österholm gratefully acknowledges financial support from Jan Wallander’s and Tom Hedelius’ Foundation.
I. INTRODUCTION

The debate on the role of money for inflation has gathered some steam recently. Woodford (2007a) and Galí et al. (2004), among others, argue that central banks should focus on interest rates rather than money when setting monetary policy. However, McCallum (2001) points out that the absence of monetary aggregates in a theoretical framework does not necessarily constitute a “non-monetary model.” Empirically, Nelson (2003) and Gerlach (2004) stress that money adds to the explanation of inflation dynamics, which should make it an important element of modern monetary policy no matter its theoretical merits.

The discussion has some importance for monetary policy. In Europe, the European Central Bank (2003) continues to stress the importance of monetary analysis in its policy framework and the Bank of England’s Mervyn King (2005, p. 18) continues to stress that “[m]oney remains at the heart of the transmission mechanism” of monetary policy, even though many central banks have decided to use interest rates as instruments. In the U.S., however, the Fed recently discontinued collecting data on certain monetary aggregates, suggesting a relative smaller emphasis on money.

At the heart of the idea of a cashless economy are current New Keynesian dynamic general equilibrium models. These models treat money mostly as a unit of account (Woodford, 2003; Galí and Gertler, 2007). Under fully separable preferences, real money demand follows standard assumptions, but money affects neither output nor inflation. The dominant channel of monetary policy transmission is the interest rate, which influences and consumption and investment decisions. To achieve control of the interest rate, central banks supply money elastically at the set rate. There is, as Woodford (2007a) states “no compelling reason to assign a prominent role to monetary aggregates in the conduct of monetary policy” in this class of models.

However, a number of authors has argued that, when it comes to money, the New Keynesian models may be based on a too specific set of assumptions. Among others, Nelson (2002), Ireland (2004), and Andrés et al. (2006) point out that allowing for non-separability of money in household utility introduces a causal link from money to output and inflation.1 In addition, Andrés et al. (2007) show that money may be a leading indicator of future economic developments in the presence of adjustment costs for real balances making money demand forward-looking. Last not least, Christiano and Rostagno (2001) and Goodfriend and McCallum (2007) extend the standard model by explicitly considering a banking sector, which also tends to elevate the role of money in the monetary policy process.

There is also little agreement on the relevance of money for monetary policy on the empirical side. As for structural evidence, Jones and Stracca (2006) find that non-separability cannot be rejected in a non-parametric approach for the euro area 1991-2005, even though there is

---

1 The effect is commonly thought to be small, however, under plausible preference specifications (McCallum, 2001). See also Woodford’s (2003) discussion of the cashless limiting economy.

Turning to non-structural approaches, there is evidence that appropriately filtered longer-run movements in money growth influence inflation, but this literature has been criticized for its reliance on smoothing the data over very long periods (OECD 2007). In addition, Nelson (2002), Ireland (2004), and Woodford (2007b) argue that, from a theoretical point of view, the long-run correlations reported are not at odds with the implications of the standard New Keynesian model. Moreover, de Grauwe and Polan (2005) have recently pointed to a possible lack of robustness of the long-run money-inflation link during low inflation episodes.

In light of this discussion, the most intriguing way forward may well be to focus on the short-run dynamics between money and inflation. A particular straightforward concept in this regard is Granger causality—that is, the question whether money improves out-of-sample forecasts of inflation. Granger causality has been dubbed the “sound and natural approach” to causality testing in a multivariate environment (Ashley et al. 1980, p. 1149). What is more, the potential ability of money to aid central bank’s ability to forecast inflation at short-term horizons is certainly highly relevant from an applied policy perspective.

In this paper, we use Bayesian autoregressive and vector-autoregressive—BAR and BVAR—models to answer the question whether money Granger-causes inflation in the United States. We find, perhaps surprisingly, that including money growth in simple VAR models of inflation does systematically improve out-of-sample forecasting accuracy. This holds for a long forecasting sample 1960-2005, as well for more recent subperiods, including the Volcker and Greenspan eras. However, the contribution of money to inflation forecasting accuracy is quantitatively limited and tends to be smaller in more recent subperiods, in particular in models that also include information on real GDP growth and interest rates.

The rest of the paper is organized as follows. Section 2 briefly reviews the relevant empirical literature. Section 3 discusses the empirical methodology and data. Sections 4 to 6 present our empirical results. Section 7 concludes.

II. Other Related Literature

There are a number of papers focusing on the relevance of monetary aggregates in simulated out-of-sample inflation forecasts, with mixed results. For the U.S., Bachmeier and Swanson

---

(2005) report that including money marginally improves inflation forecasts compared to simple AR models at certain for horizons exceeding one year; and Bachmeier et al. (2007) find marginal improvements from using money in inflation forecasts in a fully non-parametrical framework. A recent note by Hale and Jordà (2007), on other hand, stresses that money has no predictive power for U.S. inflation at any horizon. Finally, D’Agostino and Surico (2007), suggest that global liquidity has substantial predictive power for U.S. inflation. As for the euro area, Nicoletti-Altimari (2001) and Hofmann (2006) report some contribution of money inflation forecasting accuracy at horizons shorter than two years in a number of different models. Following Hofmann (2006), Scharmagl and Schumacher (2007) find a combination of monetary models or indicators produces particularly good inflation forecast. In addition, Assenmacher-Wesche and Gerlach (2006) find some evidence of within-sample Granger-causality between money and inflation in the euro area, in particular at frequencies lower than 5 years.

A relevant discussion in this context is how to estimate the autoregressive (AR) and vector-autoregressive (VAR) models dominating the recent empirical literature. Adolfson et al. (2007) argue that conventional VAR approaches, while providing for a good description of the data generating process, suffer from over-parameterization and that the large number of parameters to be estimated tends to deteriorate forecasting performance. A way out are Bayesian AR and VAR models, which use priors to shrink parameters (Doan et al., 1984; Litterman, 1986). We will use BARs and BVARs as forecasting tool in this paper.

III. METHODOLOGY AND DATA

For money to be Granger-causal for inflation we require that the forecasting performance of a BVAR model including money exceeds that of an otherwise identical BVAR or BAR model excluding money. The criterion for forecasting performance is the root mean squared error \( (RMSE) \) of the forecast.\(^5\)


\(^4\) The interpretation of these findings is complicated somewhat by the combination of a large number of overlapping approaches, including, for instance, various monetary aggregates, the so-called p-star model, trend variables, and low-frequency indicators.

\(^5\) The Diebold and Mariano (1995) test is often used to test whether the difference in forecasting performance between two models is statistically significant. However, this test is not valid in our setting with recursively generated forecasts from nested models and a forecast horizon that is larger than one (Clark and McCracken, 2001). In fact, as pointed out by Clark and McCracken (2005), no formal test presently exists for this setting. There is also the argument that in forecasting significance tests are of little value in addition to the \( RMSE \) criterion (Armstrong, 2007).
Establishing Granger-causality out-of-sample based on forecasting performance has a number of advantages. Out-of-sample tests are close to Granger’s original idea of causality testing (Ashley et al., 1980), and within-sample tests can be difficult to implement in a multivariate framework. The ability of money to improve actual inflation forecasting is also of interest to policy makers. Lastly, out-of-sample forecasting performance presents a higher hurdle for usually well-parameterized VAR models than within-sample exercises.

Our forecasting model is given by

\[ G(L)x_t = \delta + \eta_t, \quad (1) \]

where \( G(L) = I - G_1 L - \ldots - G_p L^p \) is a lag polynomial of order \( p \), \( x_t \) is an \( nx1 \) vector of stationary macroeconomic variables and \( \eta_t \) is an \( nx1 \) vector of iid error terms fulfilling \( E(\eta_t) = 0 \) and \( E(\eta_t \eta_t') = \Sigma \). In the four-variate BVAR model, we would have

\[ x_t = (\Delta p_t, \Delta m_t, \Delta y_t, i_t)', \] with \( \Delta p_t, \Delta m_t, \Delta y_t, \) and \( i_t \) representing inflation, money growth, real GDP growth, and interest rates, respectively. In the trivariate, bivariate, and univariate models, we have \( x_t = (\Delta p_t, \Delta y_t, i_t)', x_t = (\Delta p_t, \Delta m_t)', \) and \( x_t = (\Delta p_t)', \) respectively. In all models, the lag length is set to \( p = 4 \). The priors on the dynamics in model (1) are given by a Minnesota-style prior: For variables in levels (first differences), the prior mean on the coefficient on the first own lag is one (zero); all other coefficients in \( G_i \) have a prior mean of zero. The prior for the covariance matrix is a mainstream diffuse prior.

Our quarterly data ranges from 1960Q1 to 2005Q3 and were provided by the Board of Governors of the Federal Reserve System. With the exception of the interest rate, all series are seasonally adjusted. Inflation is based on the consumer price index and the monetary aggregate is M2. The interest rate is the three-month treasury bill rate.

All empirical models are estimated using growth rates computed as logarithmic first-differences (that is, quarter-on-quarter growth rates) of the original series in levels, and all variables are measured in percent. For ease of interpretation, however, we will present RMSEs converted into changes over the same quarter last year (that is, annual growth rates) in percent. Figure 1 presents the raw data in the form of annual growth rates.

---

6 Within-sample, Granger causality can be inferred from the significance of coefficients or impulse response functions (Lütkepohl, 1989; Hamilton, 1994).

7 The fact that our simulated (or “pseudo”) out-of-sample exercise is based on revised instead of real time data should have no bearing on the results. Indeed, in a recent paper, Faust and Wright (2007) argue that the use of real time or revised data does not influence the relative performance of various model- and expert-based forecasts for inflation and output.
Figure 1. Data.

IV. MONEY AND INFLATION SINCE 1960

We start by comparing the out-of-sample forecasting performance of the model in equation (1) for the full post-war period. That is, we ask whether the bivariate model produces better inflation forecasts than and the univariate model and, in addition, whether the inflation forecasting ability of the four-variate model exceeds that of the trivariate model.

The out-of-sample forecasting exercise is straightforward: All models are initially estimated for a ten-year training period from 1960Q1 to 1969Q4. We then generate dynamic forecasts for

Note: Growth rates are given as annual growth rates (that is, changes with respect to the same quarter in the preceding year). All estimations use quarter-on-quarter growth rates (that is, first differences).
the first evaluation period 1968Q1 to 1970Q4, that is, up to twelve quarters ahead. Then the sample is extended one period, the models are re-estimated, and new forecasts values up to twelve periods ahead are generated. This continues until we reach the end of our sample 2005Q2. The procedure generates between 143 and 132 inflation forecasts to evaluate depending on the forecast horizon.

The forecasts themselves are generated in a standard fashion. For every draw from the posterior distribution of the coefficients of a given model, a sequence of shocks are drawn reflecting the posterior distribution of the covariance matrix and used to generate future data. We hence get as many paths for each variable as we have iterations in the Gibbs sampling algorithm—that is, 10,000. The evaluation is then conducted using the median forecast from this generated predictive density.

Figure 2 shows the results in the form of the difference between the univariate and the bivariate RMSEs (left panel) as well as the difference between the trivariate and the fourvariate RMSEs (right panel) for the post-1960 period. As discussed, the RMSEs have been converted into annual growth rates.

Figure 2. Reduction in RMSE from adding money growth at different forecasting horizons, 1960-2005.

There are a number of noteworthy results. First, and perhaps most strikingly, the inclusion of money growth seems to clearly improve the forecasting performance during the full post-war period. We find a positive difference in RMSEs between the models excluding and including money growth, suggesting that indeed money growth Granger-causes inflation. This holds at all forecasting horizons except the first quarter, where the difference in RMSEs is estimated at zero.

Notes: Shown are RMSEs for annual growth rates.

---

9 Note that we restrict the sampling to draws which ensure the stationarity of the estimated models. This additional step does not change our results qualitatively and, as a rule, has negligible quantitative effects. (An exception are the full sample (1960-2005) RMSEs of the tri- and four-variate models, which both are approximately 0.7 percentage point higher if non-stationary draws are not rejected.)
Second, in quantitative terms, the improvement in forecasting performance from adding money growth over the full sample period are generally small. The largest reductions in \( \text{RMSEs} \) of about 0.09 percentage points occur at longer forecasting horizons.

Finally, over the full sample period the marginal contribution of money growth is comparable across models—that is, the \( \text{RMSE} \) reduction resulting from adding money to the univariate BAR model is similar in magnitude to the reduction achieved by adding money growth to the trivariate BVAR model including real GDP growth and the interest rate.

The results have potentially interesting implications. While structural interpretations of time series evidence are obviously difficult, Figure 2 does not contradict the idea that that money may have some (however small) causal or leading role in the determination of inflation.\(^{10}\) And in empirical terms, these findings seem supportive of the view that money, while perhaps useful in the euro area, is generally not helpful in predicting inflation in the U.S. with its vastly different financial sector organization and consumer habits. The question is, however, how robust these results are to changes in the time period under consideration.

V. The Declining Role of Money During the Volcker and Greenspan Period

There are reasons to suspect that the relationship between money and inflation may have changed over the years. In particular, Paul Volcker’s tenure as the chairman of the Board of Governors of the Federal Reserve System, from August 1979 to August 1987, is often mentioned as a watershed between the period of high inflation following the first oil crisis and the end of the Bretton Woods system and the period of more moderate inflation that followed (see Figure 1).\(^{11}\) If the “great moderation” changed the relation of money and inflation, we may expect the dynamic relation between money growth on inflation to change as well.

Figure 3 presents evidence for recent subperiods along the lines of the full sample exercise. The solid lines (“Post-1987”) report the improvement (or deterioration) in forecasting accuracy from adding money growth for a forecasting period starting 1987Q3—that is, post-Volcker—up to the end of our sample in 2005Q2. The short-dotted lines (“Post-1989”) are based on the forecasting period 1989Q4-2005Q2. Finally, the long-dotted lines (“Post-1993”) show the results for the forecasting period 1993Q4-2005Q2. We estimate all four models

---

\(^{10}\) One of the many problems of interpreting non-structural results is that the monetary interest rate rule may introduce links between one economic variable and another that is absent in the economic structure \textit{per se}. For instance, the empirical link money growth and inflation could be due to the Federal Reserve actively targeting a certain money growth rate. Of course, whether this is a relevant scenario is another question. A related issue raised, for instance, by Orphanides (2001, 2003) is that money may have leading qualities simply because central banks extract otherwise unavailable real time information on output movements from it. Then again, the empirical impact of money on inflation would be through the interest rate rule, making it hard to explain why empirically money growth adds to inflation forecast in addition to the interest rate.

\(^{11}\) See Goodfriend and King (2005) for a recent discussion.
discussed in the previous section for each of these subperiods, using again ten-year initial training periods.\textsuperscript{12}

\textbf{Figure 3. Reduction in RMSE from adding money growth at different forecasting horizons, 1987/89/93-2005.}

\textit{x}_t = (\Delta \rho, \Delta m_t) \text{ versus } \textit{x}_t = (\Delta \rho_t)

\textit{x}_t = (\Delta \rho, \Delta m_t, \Delta y_t, i_t) \text{ versus } \textit{x}_t = (\Delta \rho, \Delta y_t, i_t)

Notes: Shown are \textit{RMSEs} for annual growth rates. All forecasting periods extent to 2005Q2.

The results suggest, that the forecasting performance of money growth changes over time, with a clear deterioration in more recent subperiods. Note, first, that both in the univariate/bivariate comparison (left panel) and the trivariate/fourvariate comparison (right panel) the reduction in \textit{RMSEs} achieved by the inclusion of money growth in the post-1987 sample clearly exceeds the reduction reported for the full sample (see Figure 2), increasing to about 0.25 and 0.15 percentage points at longer horizons, respectively. However, the marginal impact of money growth is much smaller, in particular at longer horizons, during the post-1989 period. Finally, in the post-1993 sample, the improvement in forecasting accuracy from adding money growth falls by an order of magnitude to levels strictly below the full sample results, turning even negative at shorter inflation horizons. In other words, money growth ceases to Granger-cause inflation at all horizons in this period.

The other message conveyed by Figure 3 is that the estimated trivariate BVAR model, which includes inflation, real GDP growth, and the three month interest rate, provides a fairly complete description of the forces driving inflation in more recent subperiods. More precisely, in quantitative terms the trivariate model leaves little to be gained from the addition of money growth in the post-1993 forecasting subperiod. While adding money growth reduces \textit{RMSEs} by about 0.15 percentage points during the post-1987 subperiod, the maximum reduction falls to about 0.05 and 0.001 during the post-1989 and post-1993 subsample, respectively. These effects are small compared to results of the univariate/bivariate comparison during the same subperiods.

\textsuperscript{12} The forecast evaluations are based on between 73 and 62 periods for the “Post-1987” period, 64 and 53 for “Post-1989”, and 48 and 37 for “Post-1993”, depending on the forecast horizon.
VI. HORSE RACE

From the perspective of policy makers, in addition to the relative performance of models with and without money, another interesting question is which of the approaches would be most useful for forecasting inflation.

To answer this question, we take a look at the absolute performance of the various inflation forecasting models. To that end, Figure 4 depicts the \textit{RMSE} levels for all four inflation models—univariate, bivariate, trivariate, and four-variate—and sample periods considered in Sections 4 and 5. Panel (a) reports the \textit{RMSE} levels for the full sample and the 1970-2005 forecasting period. Panel (b) does the same for the 1987-2005 subperiod, panel (c) for 1989-2005, and panel (d) for 1993-2005.

A first result emerging is that the \textit{RMSE} levels starkly differ between the full sample and the subperiods, which exclude the 1970s and most of the 1980s. While the average \textit{RMSE} level at medium and long horizons is between 2 and 3.7 percent for the full sample across all models, the average is at or below one percent in the more recent subperiods. This impressive improvement in overall inflation forecasting accuracy seems to be a consequence of the reduction in inflation levels and volatility during the “great moderation” in the early 1980s—predicting price changes and avoiding large errors obviously is easier in a low-inflation environment.\footnote{The moderation influenced the behavior of money growth and the other times series as well (see Figure 1). Among other things, this lead to an overall change in the dynamics underlying the estimated models. This is illustrated by Figure A1 in the Appendix, which shows the impulse response functions of the simple bivariate BVAR model of inflation and money growth for the pre-Volcker years (1960-1979) and the Volcker-Greenspan years (1980-2005). A striking difference between both periods is the larger and more persistent impact of shocks 1960-1979.}

As to model selection, Figure 4 suggest that the choice between the bivariate and four-variate money-based BVAR models is a matter of the forecasting horizon. For the full sample period, the bivariate outperforms the four-variate BVAR. In contrast, in the more recent subperiods, the four-variate BVAR model dominates the bivariate BVAR. In general, the differences in forecasting accuracy at shorter horizons remains small, however.
VII. CONCLUSION

We find strong indications that including money growth Granger-causes inflation in the United States, that is, it improves forecasting performance out-of-sample. For forecasting horizons up to 12 quarters ahead, bivariate BVAR models of inflation and money growth and fourvariate BVAR models, including, in addition, real GDP growth and interest rates, consistently produce better inflation forecasts than models excluding money.
These results may be surprising given that current monetary theory often eschews monetary aggregates as a leading or causal factor for inflation. While any structural interpretation of time series evidence can be only suggestive, the finding that money growth Granger-causes inflation does support (and certainly does not contradict) the idea that money may have some, however small, role in the determination of inflation in the shorter run. This seems to go against the general idea of many current monetary models, where money is either excluded or strictly endogenous with regard to, in particular, interest rates and output. Thus, one way to interpret these findings is that there may be room for more general approaches taking into account, for example, financial frictions or portfolio adjustment costs.

It is important to keep in mind, however, that the quantitative role played by money in forecasting inflation is limited. Including money growth, while improving forecasting accuracy as such, does not yield dramatically better predictions—in particular during more recent subperiods and when using fuller models. At least from a practical perspective, this suggests that the opportunity costs of ignoring money, while positive, may be small.
Figure A1. Impulse response functions from bivariate BVAR models including money growth and inflation, 1960-1979 and 1980-2005.

(a) Pre-Volcker Period (1960-1979)

(b) Volcker-Greenspan Period (1980-2005)

Notes: Horizon in quarters. Colored band is 68 percent confidence band. The Figures show the impulse response functions of the simple bivariate BVAR model of inflation and money growth for the pre-Volcker years (1960-1979) and the Volcker-Greenspan years (1980-2005). The impulse responses were calculated using a standard Cholesky decomposition of the covariance matrix, where CPI inflation was ordered ahead of money growth.
REFERENCES


Woodford, Michael (2007a), How Important is Money in the Conduct of Monetary Policy, *NBER Working Paper* 13325.