The Quantity Theory Revisited:
A New Structural Approach

Makram El-Shagi

Sebastian Giesen

Logan J. Kelly

The Quantity Theory Revisited:
A New Structural Approach

Makram El-Shagi

Sebastian Giesen

Logan J. Kelly

April 2011   No. 7

Authors: *Makram El-Shagi*
             Department Macroeconomics
             E-mail: Makram.El-Shagi@iwh-halle.de
             Phone: +49 (0) 345 7753-835

             Sebastian Giesen
             Department Macroeconomics
             E-mail: Sebastian.Giesen@iwh-halle.de
             Phone: +49 (0) 345 7753-804

             Logan J. Kelly
             University of Wisconsin-River Falls
             E-mail: Logan.Kelly@uwrf.edu

The Quantity Theory Revisited:
A New Structural Approach[*]

Abstract

While the long run relation between money and inflation is well established, empirical evidence on the adjustment to the long run equilibrium is very heterogeneous. In this paper, we show that the development of US consumer price inflation between 1960Q1 and 2005Q4 is strongly driven by money overhang. To this end, we use a multivariate state space framework that substantially expands the traditional vector error correction approach. This approach allows us to estimate the persistent components of velocity and GDP. A sign restriction approach is subsequently used to identify the structural shocks to the signal equations of the state space model that explain money growth, inflation, and GDP growth. We also account for the possibility that measurement error exhibited by simple-sum monetary aggregates causes the consequences of monetary shocks to be improperly identified by using a Divisia monetary aggregate. Our findings suggest that when the money is measured using a reputable index number, the quantity theory holds for the United States.

Keywords: Divisia money, state space decomposition, sign restrictions

JEL Classification: E31; E52; C32

Die Quantitätstheorie:
Ein neuer struktureller Ansatz

Zusammenfassung

Im Gegensatz zur weitgehend etablierten Langfristbeziehung zwischen Geldmenge und Preisen ist die empirische Evidenz bezüglich des Anpassungsprozesses hin zu diesem Langfristgleichgewicht durchwachsen. Der vorliegende Beitrag untersucht diesen Zusammenhang für die USA im Zeitraum von 1960 bis 2005. Dazu wird ein multivariates Zustandsraummodell verwendet, das den herkömmlichen Kointegrationsansatz, mit dem diese Frage im Regelfall analysiert wird, deutlich erweitert. Dieser Ansatz erlaubt die Identifikation der persistenten Komponente der Geldumlaufsgeschwindigkeit und des Bruttoinlandsprodukts. Aufbauend auf dem Zustandsraummodell erfolgt über Vorzeichenrestriktionen die Identifikation struktureller Schocks auf die "Signalgleichungen", welche die Entwicklung von Geldmengenwachstum, Inflation und Wachstum beschreiben. Durch die Verwendung eines so genannten Divisa Geldmengenaggregats wird gleichzeitig der Möglichkeit Rechnung getragen, dass der Zusammenhang zwischen Geld und Preisen durch Messfehler einfacher Geldmengenaggregate verzerrt wird. Die Resultate zeigen, dass ein quantitätstheoretischer Zusammenhang im Sinne eines maßgeblichen Zusammenhangs von Geld und Preisen bei Verwendung eines angemessenen Geldmengenaggregats klar erkennbar ist.

Schlagwörter: Divisia Geld, Zustandsraummodell, Vorzeichenrestriktionen

JEL-Klassifikation: E31; E52; C32

1. Introduction

While the quantity theory has often been validated using cross country data (Teles and Uhlig 2010, Herwartz and Reimers 2006, McCandless and Weber 1995), tracking the transmission of monetary shocks within a given economy remains one of the most hotly debated topics of the profession. In this paper, we attempt to combine the three main strands of literature dealing with this issue: (i) the P-Star literature, (ii) the structural VAR literature and (iii) the statistical index number literature.

P-Star literature   After interest rates replaced monetary aggregates as the central driving force of inflation in applied macroeconomics models, the P-Star literature strived to reestablish the role of money supply in explaining price movements as indicated by the quantity theory (Hallman, Porter, and Small 1991). The most widespread version of these models follows the two-pillar-Phillips curve approach proposed by Gerlach and Svensson (2003), where a standard Phillips-Curve framework is extended by including a measure of excess liquidity. While these models have been applied to the United States, they have been particularly successful in explaining inflation in Europe, especially when long samples are used, see e.g. Kaufmann and Kugler (2008), Svensson (2000), Tödter (2002), Tödter and Reimers (1994). Movements of the equilibrium of velocity, however, have proven to be a key obstacle to the identification of excess liquidity, which is essential to estimating these models. Most approaches that try to include nonlinear movements of velocity rely on explaining money demand either through the opportunity costs of holding money (Orphanides and Porter 2000) or by accounting for the role of wealth and asset prices (Dreger and Wolters 2009).

In this paper, we follow an alternative P-star approach proposed by El-Shagi and Giesen (2010) and El-Shagi (2010), who consider any persistent movement to reflect a change of equilibria. To do so, El-Shagi and Giesen replace the standard cointegration techniques that are commonly applied to estimate P-star models with a state space framework, where equilibrium velocity and potential output are treated as latent variables following a random walk with drift. While this approach does not allow us to analyze the driving forces of velocity, we are able to ensure that equilibrium movements due to changes in unobserved indicators are captured.

Structural VAR literature   The P-star literature deals mostly with the question of whether excess liquidity causes inflation. The structural VAR literature, however, deals with the identification of monetary shocks and transmission of those shocks throughout the economy, e.g. Sims (1980). Most of the early literature applies standard SVAR, where zero restrictions are used to identify structural shocks.[1] In the last decade, sign restrictions have become a popular method of identifying structural shocks (Faust 1998, Canova and De Nicoló 2002, Uhlig 2005). The motivation for using sign restrictions, at least partly, is to overcome the empirical "puzzles" that are found in most econometric models that include both money and interest rates. However, they generally allow us to embed more convincing economic explanations into structural econometric models than do other structural identifications, such as the standard Cholesky decomposition.

Statistical index number literature   We account for the possibility that the failure to properly identify the consequences of monetary shocks is caused by measurement error exhibited by simple-sum monetary aggregates. Simple-sum monetary aggregates are commonly reported as the official monetary aggregates by central banks throughout the world. Barnett (1980) and others have long argued that the use of simple-sum monetary aggregates is theoretically indefensible in a modern economy.[2]

The key problem is that simple-sum aggregation methodology fails to account for the dual nature of monetary assets. Most monetary assets provide simultaneously both monetary and investment services. Simple-sum aggregates (i) treat all monetary aggregates as though they provide equal levels of monetary service, and (ii) ignore investment service flow entirely. As a result, simple-sum aggregates have no theoretical link to the monetary services flow, significantly over state the money stock, and mask the dynamics of the relationship between money and interest rates.

These problems cannot be avoided by simply narrowing the definition of money to include fewer assets as this arbitrarily ignores many assets that provide monetary service. These problems can only be addressed by carefully measuring money using

_____

[1] This includes the seminal contributions of Gordon and Leeper (1994), Leeper, Sims, and Zha (1996) Bernanke and Mihov (1998), Christiano, Eichenbaum, and Evans (1996) and Christiano, Eichenbaum, and Evans (1999).

[2] See, e.g., Barnett and Serletis (2000), Barnett, Chae, and Keating (2005), Barnett, Keating, and Kelly (2008), Barnett and Chauvet (2011), Kelly (2009), and Kelly, Barnett, and Keating (2011).

reputable index numbers based on assumptions that are internally consistent with the models in which the monetary data will be used. The Divisia Monetary Index (Divisia money) is one such index number, and is published by the Federal Reserve Bank of St. Louis as the Monetary Service Index (MSI) for the United States.

The contribution of this paper is to propose a unified identification scheme that allows us to identify monetary shocks and track their propagation through the economy using a parsimonious specification that includes only (Divisa) money, prices and GDP. To do this, we adjust the state space framework used in El-Shagi and Giesen (2010) and El-Shagi (2010) to account for contemporaneous impacts of shocks to the latent variables and apply a sign restriction identification scheme to the remaining reduced form given by the signal equations. To make sure that our results are not distorted by the measurement error exhibited by the official United States monetary data, we employ the Monetary Service Index (MSI), which is a Divisia monetary aggregate.[3] Thus, we combine three of the most notable approaches dealing with the consequences of money supply shocks. This allows us to show that the quantity theory of money is still relevant to price movements and that - contrary to its intent - monetary policy in the US did not diminish business cycles, but instead substantially contributed to fluctuations of both prices and GDP.

The remainder of the paper is structured as follows. Since the use of Divisa money is quite unusual in this literature, sections two and three will start by introducing the dataset and giving a brief descriptive analysis of the core developments. Section four outlines some elemental theoretical considerations that underly our econometric specification and provides an overview of the identification strategy. Section five describes the state space framework. The identification scheme for the remaining structural shocks is given in section six. The second paragraph of this section presents the results, including both the parameter estimates and the impulse responses. Some summarizing comments and the conclusions are found in section seven.

2. Divisia Money

The official monetary aggregates   The official monetary aggregates released by the United States Federal Reserve are calculated using simple-sum methodology.

---

[3] Technically, the MSI is calculated using the Tornqvist-Theil index number, which is the discret-time approximation of the continuous-time Divisia index number.

Simple-sum methodology involves simply summing the nominal values of all monetary assets in a giving grouping of assets. This method, at first look, is an intuitive way to measure the nominal money stock, since the nominal price of a portfolio of monetary assets is the simple-sum. Barnett (1980), however, showed that the use of simple-sum aggregates can only be justified if all monetary assets, including currency, yield the same own rate of interest. Moreover, simple-sum methodology ignores the fact that an interest yielding monetary asset is a joint product of a monetary and an investment asset. Barnett, Chae, and Keating (2005) and Barnett, Keating, and Kelly (2008) argue that because of this confounding, the official monetary aggregates confound the money stock with the present value of the investment yield returned by monetary assets, and thus significantly over-state the money stock. Kelly (2009) showed that this confounding causes simple-sum aggregates to fail to capture the true relationship between the economic money stock and interest rates. Finally, Kelly, Barnett, and Keating (2011) argues that much of the liquidity puzzle can be explained by the measurement error exhibited by simple-sum aggregation.

Understanding Divisia monetary aggregates Given that the use of Divisia money is unusual in this literature, a brief review of the derivation of the Divisia index number first derived by Francois Divisia in 1925 and its application to monetary aggregation may be useful. Let the representative consumer's current period inter-temporal utility function, u_t, be weakly separable in each period's consumption of goods and monetary assets, so that u_t can be written

$$
\begin{aligned}
u_t &= u_t\left(\nu_t, \nu_{t+1}, \ldots, \nu_{t+T}; c_t, c_{t+1}, \ldots, c_{t+T}, B_{t+T}\right) \qquad (1)\\
&= \mathcal{U}_t\left[M\left(\nu_t\right), M_{t+1}\left(\nu_{t+1}\right), \ldots, M_{t+T}\left(\nu_{t+T}\right);\right.\\
&\qquad \left. V\left(c_t\right), V_{t+1}\left(c_{t+1}\right), \ldots, V_{t+T}\left(c_{t+T}\right), B_{t+T}\right],
\end{aligned}
$$

where ν_t is a vector describing the monetary portfolio in period t, c_t is consumption in period t and B_{t+T} is the quantity of the benchmark asset held at the end of the planning horizon. Note that $M\left(\nu_t\right)$ is the representative consumer's category utility function over monetary assets. Assuming $M\left(\nu_t\right)$ is linearly homogeneous, then $M\left(\nu_t\right)$ is also the economic quantity aggregate.[4] Following Barnett (1980),

───

[4] The assumption of linear homogeneity is not necessary for the existence of a quantity aggregator function; however, in the absence of linear homogeneity, the category utility function cannot be interpreted as the quantity aggregator function. For further explanation of aggregation without linear homogeneity, see Barnett and Serletis (2000).

totally differentiate $M\left(\nu_t\right)$ to obtain

$$dM\left(\nu_t\right) = \sum_{n=1}^{N} \frac{\partial M}{\partial \nu_{nt}} d\nu_{nt}. \tag{2}$$

Using the first order conditions of the representative consumer's problem, we replace $\frac{\partial M}{\partial \nu_{nt}}$, which each contain unobserved parameters, with $\lambda\psi_{nt}$, where λ is the Lagrange multiplier and ψ_{nt} is the user cost of asset n in period t. This substitution yields

$$\frac{dM\left(\nu_t\right)}{\lambda} = \sum_{n=1}^{N} \psi_{nt} d\nu_{nt}. \tag{3}$$

Barnett (1980) showed that $\lambda = \Psi\left(\psi_t\right)^{-1}$, where $\Psi\left(\psi_t\right)$ is the dual price index of money. Substituting into 3 yields

$$\Psi\left(\psi\right) dM\left(\nu_t\right) = \sum_{n=1}^{N} \psi_{nt} d\nu_{nt}. \tag{4}$$

Converting 4 to growth rate form yields

$$d\log M\left(\nu_t\right) = \sum_{n=1}^{N} \omega_{nt} d\log \nu_{nt}, \tag{5}$$

where ω_{nt} is the expenditure share on monetary asset n in period t. Equation 5 is the Divisia quantity monetary aggregate in growth rate terms. Note that 5 is a differential equation. We find the level quantity aggregate by solving 5,

$$M\left(t\right) = \exp \int_{\tau=0}^{\tau=t} \left[\sum_{n=1}^{N} \omega_{nt}\left(\tau\right) \frac{d\log \nu_{nt}\left(\tau\right)}{d\tau} \right] d\tau, \tag{6}$$

where $\omega_{nt}\left(t\right)$ is the time path of expenditure shares on monetary asset n and $\nu_{nt}\left(t\right)$ is the time path of asset quantities. In discrete time, we use the Tornquest-Theil approximation, which is just the Simpson's rule approximation of 6

$$\log M\left(\nu_t\right) - \log M\left(\nu_{t-1}\right) = \sum_{n=1}^{N} \bar{\omega}_{nt} \left(\log \nu_{nt} - \log \nu_{n,t-1}\right), \tag{7}$$

where

$$\bar{\omega}_{nt} = \frac{\psi_{nt} + \psi_{n,t-1}}{2}.$$

Thus, we can see that the Divisia monetary aggregate measures the monetary service flow yielded by the current monetary portfolio.

3. Data and Descriptive Analysis

For our analysis we use quarterly US data series. Our sample includes data from 1960Q1 to 2005Q4. The Monetary Service Index, the Divisa monetary aggregate produced by the Federal Reserve Bank of St Louis, is only available during this period. All data series are seasonally adjusted. We denote the price indicator p to represent the Consumer Price Index (CPI), which we obtained from the Bureau of Labor Statistics. Production, y, is defined as Gross Domestic Product (GDP) throughout the paper and is obtained from the Bureau of Economic Analysis. Following Barnett, Offenbacher, and Spindt (1984) and Kelly (2009) we use the Monetary Service Index at the M3 level of aggregation (MSIM3) as the monetary aggregate, m, which is obtained from the Federal Reserve Bank of St. Louis. Thus, the vector of interest is $x = [m, p, y]'$.

Conventional wisdom suggests that using a sufficiently narrow monetary aggregate is necessary to avoid the many empirical puzzles that have plagued monetary economists; however, Kelly, Barnett, and Keating (2011) finds that when money is measured using a reputable index number, a broad monetary aggregate exhibits empirical results that are more consistent with theoretical expectations than do narrow aggregates. Specifically, the liquidity puzzle found by Leeper and Gordon (1992) and others is largely resolved when a Divisia M3 is used to measure money. Moreover, using a reputable index number avoids arbitrarily discarding information.

The data appendix presents an overview of the individual data series in logarithms (See Figure A.1). All data series are identified to be upward trending. Looking at the first differences we see the series to be affected by some persistent behavior, especially for the CPI variable. The log difference of each series is tested for stationarity using the Dickey-Fuller (DF) and Ng-Perron (NP) unit root tests (Ng and Perron 2001).

Both the DF and NP tests indicate that the log difference of MSIM3 and GDP are stationary at one percent significance level. The DF and NP tests indicate that the log difference of CPI is stationary at the ten and five percent significance level,

Table 1: Dickey-Fuller Tests

Dickey-Fuller	MSIM3	GDP	CPI
Test Statistic	-5.58	-10.06	-2.73
P-value	0.00	0.00	0.07

respectively (see Tables 1 and 2). Therefore, we will model each series as stationary in log differences.

Table 2: Ng-Perron Unit Root Tests

Ng-Perron	MZa	MZt	MSB	MPT
Test Statistic for MSIM3	-27.68	-3.69	0.13	0.97
1%	-13.80	-2.58	0.17	1.78
Test Statistic for GDP	-21.63	-3.28	0.15	1.17
1%	-13.80	-2.58	0.17	1.78
Test Statistic for CPI	-11.08	-2.34	0.21	2.26
1%	-13.80	-2.58	0.17	1.78
5%	-8.10	-1.98	0.23	3.17

4. Theoretical considerations and empirical strategy

The core model   Our model aims at explaining money, prices and output. While we allow all three variables - and the derived disequilibria - to affect the others as in a standard VAR setup, the individual equations quite closely mirror widespread theoretical concepts.

Price movements are captured by a two-pillar Phillips curve. That is, inflation is driven by real indicators (in our case these are the output gap and GDP growth) and monetary indicators, i.e. excess liquidity and money growth. Further, we account for the momentum of inflation dynamics, i.e. we include an autoregressive process that has been established by the previous literature.

Although we employ a monetary quantity aggregate instead of using an interest rate as a monetary indicator, our model of money growth is inspired by a Taylor-rule specification, i.e. we include inflation dynamics and real dynamics as driving forces, to capture the major objectives of monetary policy. However, the coefficients cannot entirely be interpreted along the lines of a Taylor rule specification since money is partly determined endogenously by market participants. We also account for the momentum of money growth. While this might reflect a smoothing preference of the central bank, it is more likely that it results from the time consuming transmission process from monetary policy actions to the monetary aggregates. Finally, we add excess liquidity to allow the central bank to compensate for its past mistakes.

GDP movements are expected to rely mostly on their own momentum and - more importantly - the attraction of potential output. Including monetary indicators allows the central bank to actually affect the business cycle. Prices are necessary to account for the joint dynamics created by supply and demand shocks.

The importance of velocity   We start from a standard log linearized quantity equation

$$m + v = y + p, \tag{8}$$

where m, v, y, and p are the natural logarithms of money, velocity, production and prices.

Suppose that to be in general equilibrium, $v = v^*$, where the star marks equilibrium values. Then an increase in money supply constitutes the creation of excess liquidity by construction. It can easily be seen that this increase has to be reflected by a corresponding decrease of velocity if both prices and production exhibit some degree of rigidity. More precisely, the negative deviation of velocity from equilibrium corresponds to excess liquidity. Accordingly, changes in the persistent component of velocity can be regarded as changes in money demand, since - given the current prices and production - higher money supply would be needed, to achieve equilibrium.

The importance of using correct aggregation methodology   The use of a considerably more complex Divisia monetary aggregate begs the question: Does monetary aggregation methodology, simple-sum or Divisia, matter? We will attempt to illustrate the importance of using correct aggregation methodology with a rudi-

mentary graphical analysis in the spirit of Rudebusch and Svensson (2002). We begin with the standard log-linearized quantity equation as before. To maintain consistency with Rudebusch and Svensson (2002), we use GDP Deflator instead of CPI in this section only. Rearranging the quantity equation yields

$$m - y = p - v. \tag{9}$$

From an examination of equation 9, we expect to find a positive, one-to-one, relationship between $m - y$ and p. However, figure 1 reveals that when simple sum monetary aggregates are used, the correlation is negative. On the other hand, when a Divisia monetary aggregate is used we see the expected positive relationship with a correlation coefficient of nearly one. Admittedly, this naive graphical analysis is hardly conclusive, but it does serve to illustrate the distortions of economic relationships that can result from using simple-sum aggregation methodology.

Figure 1: Log of the ratio of the monetary aggregate and real gross domestic product vs. the log of the GDP Deflator price index.
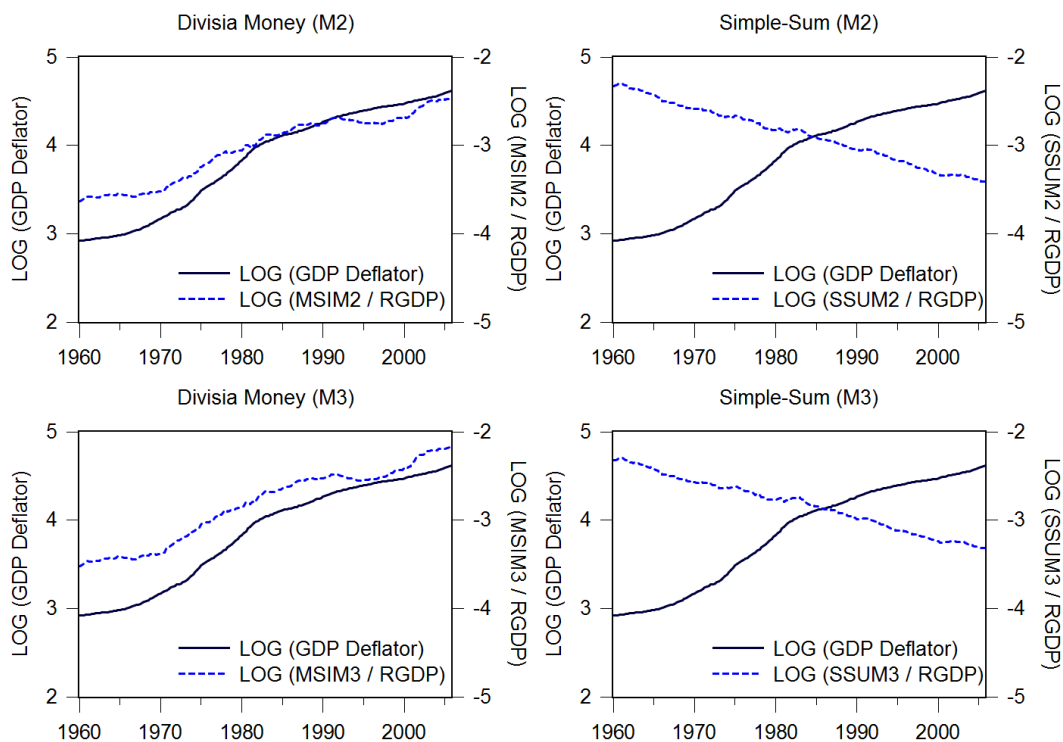
Business cycle correction Decomposing total velocity to obtain an estimate of excess liquidity may create an undesirable statistical artifact, as shown by El-Shagi (2010). In a recession, production, y, falls below potential output, y^*. Since neither money nor prices can plausibly compensate for this contemporaneously, the increased output gap is once again mirrored by a corresponding deviation of velocity from equilibrium. Strictly speaking, this constitutes excess liquidity, since less money is needed during the recession. However, the return of velocity to equilibrium takes place in the course of output returning to potential output, which may create a spatial correlation between "excess liquidity" and output growth. Thus, we define excess liquidity such that it does not rely on the decomposition of total velocity but on the decomposition of an adjusted velocity index v^a:

$$m - p - y^* = -v^a = -(\tilde{v}^a + v^{a*}), \tag{10}$$

where the tilde marks the cyclical component. For simplicity we drop the index "a" in the components of velocity. The variables $\tilde{v}$ and v^* refer to the components of adjusted velocity in the remainder of the paper, unless noted otherwise.

Empirical strategy We propose a two step identification scheme for structural shocks that combines the two most established approaches to structural shock identification: (i) state space decomposition (using a Kalman filter) and (ii) structural VAR (SVAR).

Step 1: We identify the persistent developments in the data, more precisely the potential output and equilibrium velocity, using a state space model. Our approach roughly follows El-Shagi and Giesen (2010) and El-Shagi (2010), who embed a P-Star type model (originally developed by Hallman, Porter, and Small (1991)) in its two-pillar Phillips curve version (Gerlach and Svensson (2003)) using a state space framework. Shocks to potential output and equilibrium velocity are modeled as latent variables and estimated using a Kalman filter. Since only supply shocks can plausibly affect output persistently, the shocks to potential output are interpreted as supply shocks. Likewise, shocks to the equilibrium velocity reflect money demand shocks as outlined above.

Step 2: The signal equations of the state space model are treated as a reduced form (first difference) VAR. Since shocks to supply and money demand have already been

removed from the data, the remaining shocks reflect the consequences of structural shocks to demand, money supply and prices. We employ a sign restriction approach for the structural identification, as proposed by Canova and De Nicoló (2002) and Uhlig (2005) with the extensions proposed by Fry and Pagan (2007).

5. Identification

Step 1: The state space model The state space model used tracks the unobservable components, potential output and equilibrium velocity, within a standard two-pillar Phillips curve framework, which takes the following form. The signal equations are given by:

$$
\begin{bmatrix} \Delta m \\ \Delta p \\ \Delta y \end{bmatrix}_t = A_1 \tilde{v}_t + A_2 \tilde{y}_t + A_3 \Delta v^* + A_4 \Delta(y^*) + B(L) \begin{bmatrix} \Delta m \\ \Delta p \\ \Delta y \end{bmatrix}_t + u_t. \tag{11}
$$

Here the A_i describe coefficient vectors corresponding to the filtered variables and $B(L)$ describes coefficient matrices corresponding to the lagged variables, respectively. The cyclical components are marked with a tilde and trend variables are marked with a star. These are more formally defined as:

$$
\tilde{v} = (m - p - y^* - (-v^*))
$$
$$
\tilde{y} = (y - y^*)
$$
$$
\Delta v^* = (-v^*) - (-v^*_{t-1}))
$$
$$
\Delta(y^*) = (y^* - y^*_{t-1}).
$$

The two state equations that describe the behavior of the components of velocity are given by:

$$
v^*_t = \mu_1 + v^*_{t-1} + \varepsilon_{1t}
$$

$$
y^*_t = \mu_2 + y^*_{t-1} + \varepsilon_{2t},
$$

where μ_i describes the drift term, and ε_{it} is an i.i.d. error term.

The signal equations do not only include the level of the states but also the first difference. Corresponding to the structural identification from the error terms of the reduced form signal equations (see step 2), this allows the identification of a contemporaneous effect of a shock to a state variable that differs from the impact that can be derived from the long run adjustment process. The states are filtered using a Kalman filter with parameter estimates obtained by numerical maximum likelihood.[5]

Contrary to the standard setup used in this type of decomposition, we only model the trend component explicitly. The cycle is only implicitly included. This ensures that the various shocks are uniquely identified. If trend and cycle were modeled separately, there would be distinct forecasts for total GDP and velocity resulting from the state equations and the signal equations. Thus, it is not clear how to choose which error term is used to identify structural shocks; however, we avoid this problem with our model setup. The common stationarity restrictions on the cyclical components are nevertheless enforced on the implicit cyclical components.

While the equation that defines the cyclical component, implies an income elasticity of one, it has been shown by El-Shagi and Giesen (2010) that the filtering process is robust to deviations. However, the preliminary finding that the income elasticity of Divisa demand is close to one (see section 4) allows us to interpret the shocks to velocity as shocks to money demand. If there was a strong deviation, the error vector ε_1 would capture both the original shock to money demand and the adjustment caused by the supply side shock.

The economic story behind this type of model becomes more obvious when rephrased into an equivalent VECM form, where the latent variables are considered to be part of the vector of endogenous variables, described by the VEC process:

_____

[5] The likelihood of a state space model is estimated using prediction errors rather than residuals from the smoothed model. Since the prediction of the change of the trend is deterministic (and given by $[\mu_1\ \mu_2]'$) a change of the corresponding parameter estimates can be offset completely by a corresponding change in the constant term without affecting forecast quality. However, a change in the regression parameter does affect the predicted variance of the forecast errors. This mismatch is captured by the likelihood function, resulting in lower likelihoods when the parameter estimated is far from the true value. This allows estimation although there is no immediate impact of the regression parameters on forecast quality.

$$
\begin{bmatrix} \Delta m \\ \Delta p \\ \Delta y \\ \Delta v^* \\ \Delta y^* \end{bmatrix}_t = \begin{bmatrix} h_{11} & h_{21} \\ h_{12} & h_{22} \\ h_{13} & h_{23} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}' \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ -1 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} m \\ p \\ y \\ v^* \\ y^* \end{bmatrix}_{t-1}
$$

$$
+ A(L) \begin{bmatrix} \Delta m \\ \Delta p \\ \Delta y \\ \Delta v^* \\ \Delta y^* \end{bmatrix}_{t-1} + \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}.
$$

The coefficient matrices $A(L)$ which correspond to the lagged variables are defined as:

$$
A(i) = \begin{bmatrix} a(i)_{11} & a(i)_{12} & a(i)_{13} & b(i)_{14} & b(i)_{15} \\ a(i)_{21} & a(i)_{22} & a(i)_{23} & b(i)_{24} & b(i)_{25} \\ a(i)_{31} & a(i)_{32} & a(i)_{33} & b(i)_{34} & b(i)_{35} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},
$$

where $b(i)_{mn} = 0$ for every lag $i > 1$.

Step 2: Sign restrictions   To identify our structural innovations, we adopt the sign restriction approach introduced by Uhlig (2005), Canova and De Nicoló (2002), and Fry and Pagan (2010). Basic to this method of shock identification is the requirement that the shocks should be uncorrelated. Recalling our model from equation 11, we obtain the structural shocks by estimating the following system:

$$
\begin{bmatrix} \Delta m \\ \Delta p \\ \Delta y \end{bmatrix}_t = A_1 \tilde{v}_t + A_2 \tilde{y}_t + A_3 \Delta v^* + A_4 \Delta(y^*) + B(L) \begin{bmatrix} \Delta m \\ \Delta p \\ \Delta y \end{bmatrix}_t + \Gamma^{-1} \xi_t, \quad (12)
$$

such that $\hat{u}_t = \hat{\Gamma}^{-1}\hat{\xi}_t$, where the structural (i.e. orthogonal) residuals are denoted by $\hat{\xi}_t$. We define a normalized residual vector $\eta_t = S^{-1}\hat{\xi}_t$, such that the diagonal elements S are the standard deviations of $\hat{\xi}$. Then we are able to write the VAR residuals as:

$$\hat{u}_t = T\eta, \tag{13}$$

where $T = \hat{\Gamma}^{-1}S$.

To identify a *set* of estimated structural shocks, Fry and Pagan (2007) define a matrix Q such that $Q'Q = QQ' = I$, where I is an identity matrix, and

$$\hat{u}_t = TQ'Q\eta_t = T^*\eta_t^*. \tag{14}$$

Valid choices for Q include any Givens matrix or product of Givens matrices. Since there is an infinite number of Givens matrices, there is an infinite set of possible T^*, that shall be denoted $\mathfrak{T}$. We can generate all possible series of structural shocks that fulfill the required assumptions from a single estimate of Γ that can be obtained by a Cholesky decomposition of the covariance matrix $\varepsilon'\varepsilon$. The covariance matrix of all η^* generated by these Givens rotations[6] of the original η is equal to the identity matrix, I, since $E(\eta_t^*\eta_t^{*\prime}) = QE(\eta_t\eta_t)Q' = I$.

Following Canova and De Nicoló and Fry and Pagan we want to restrict the economic analysis to a subset of $\mathfrak{T}$ where the initial, reduced form response to any structural shocks follows some economic reasoning.[7]

We will only consider T^* where the following sign restrictions hold:

$$\begin{bmatrix} + & * & - \\ + & + & + \\ + & - & + \end{bmatrix}. \tag{15}$$

_____

[6] By Givens rotation, we mean multiplication by one or more Givens matrices. For more details see section 3.2 in Fry and Pagan (2007).

[7] This differs slightly from Uhlig (2005) who remains "agnostic" about the initial response of the monetary shock.

This implies that a money supply shock has positive contemporaneous effect on money itself, prices and GDP. Thus, we allow for an immediate inflationary response to a money supply shock and the possibility that money is non-neutral in the short-run.

Prices have a positive impact on prices and a negative impact on real GDP; as in Leeper, Sims, and Zha (1996). A negative contemporaneous effect on real GDP is plausible since an exogenous shock to prices, such as the oil price shocks of the 1970's for example, should lead to a reduction in short-run equilibrium output. Many models treat these shocks as supply side effects, and most likely the supply side dominates the transmission of these shocks throughout the real economy. We nevertheless refer to these shocks as price shocks, to clearly distinguish them from the supply shocks caused by shifts in the production function as identified by the state space approach.

We remain agnostic about the impact of prices on money. For if the exogenous prices shock is mistaken for inflation caused by excessive money growth, the central bank might react by reducing money supply. However, if the central bank recognizes the price shock to be exogenous, they might be willing to accommodate the rising demand for money in order to prevent contractionary "second round effects," i.e increasing interest rates. Also, since money is partly endogenous in the short run, the banks might be willing to provide more money under the same monetary policy.

Finally, demand has a positive impact on GDP and prices and a negative impact on money supply. A demand shock has a negative impact on money supply because of the widespread implementation of counter cyclical monetary policy. We further assume that the largest shock to any reduced form equation comes from the corresponding shock, i.e. the strongest shock to money comes from money supply, etc.

To compute our impulse responses and the corresponding confidence bounds, we combine a moving blocks bootstrap of the coefficients of the state space model with randomly generated Givens rotations. This allows us to account for the uncertainty from the estimation of the state space model and the uncertainty about T^* simultaneously. We bootstrap 2,000 estimates for the coefficient matrices A and B and the corresponding covariance matrices $\hat{u}'\hat{u}$. We then draw from these coefficient estimates and randomly generate Givens rotations until we generate a matrix, T^*, where the sign restrictions (15) hold. If this is not achieved within 200,000 draws the corresponding coefficient matrices are discarded. This procedure is repeated until we obtain 2,000 sets $\{A, B, T^*\}$. To compute the point estimates of the impulse responses, we follow Fry and Pagan. The impulse responses we report are selected

from our bootstraps by minimizing the squared differences between the impulse response and the median of the confidence bounds. Fry and Pagan argue that this is a more appropriate approach than the commonly used mean of the confidence bounds, since the mean of the confidence bounds is not necessarily a possible response that could be generated by a single model.

6. Results

Specification tests To determine the optimal lag order, p, of our dynamic model, we use the standard battery of model selection procedures. We consider models with up to five lags. All criteria employed - namely the Akaike information criterion (AIC), the Bayesian or Schwarz information criterion (BIC), and the Hannan-Quinn criterion (HQ) - indicate a lag length of order $\hat{p} = 3$. The full test results are reported in Table 3.

Table 3: Information Criteria

Lags	AIC		BIC		HQ	
1	-19.10		-18.94		-19.03	
2	-21.20		-20.88		-21.07	
3	-23.58	*	-23.11	*	-23.39	*
4	-22.71		-22.07		-22.45	
5	-19.46		-18.67		-19.14	

Notes: The asterisk marks the estimated lag order.

While most techniques that test for autocorrelation in VAR models (Portmanteau, Breusch-Godfrey) can cope with exogenous variables, it is not quite clear from a theoretical perspective how the latent variables should be treated. Thus, we test for serial correlation using separate tests to check the autocorrelation and cross correlation of residuals for up to 8 lags. To test whether single significant results are meaningful or just a consequence of the multiple testing problem, we combine the p statistics of the individual tests using a Fischer test (see, e.g. Fisher (1932)).[8] We find significant autocorrelation in the second lag - with a test statistic value of 62.5,

---

[8] This test has recently regained popularity since it was proposed by Maddala and Wu (1999) to combine time series unit root tests to serve as a panel unit root test.

which corresponds to a chi-squared probability value of 0.004 - that cannot be fully removed by adjusting the lag structure. Thus, the reported significance levels are based on heteroscedasticity and autocorrelation consistent standard errors (Newey and West 1987). The confidence bounds for the impulse responses are obtained by a moving block bootstrap (with a block length of 7) that has been shown to be robust to autocorrelation (Li and Maddala 1996, Davidson and MacKinnon 2006). We double check the parameter estimates with estimates obtained from this bootstrap and find them to be roughly identical, as far as significant impacts are concerned.

Coefficient estimates The coefficient estimates mostly confirm our hypotheses. See table 4.

Table 4: Coefficient Estimates

|	Δm		Δp		Δy	
$\tilde{v}$	-0.025	**	0.024	**	-0.010	
$\tilde{y}$	0.037		0.010		-0.085	**
Δv^*	-0.025		-0.085		-0.037	
Δy^*	0.058		0.063		-1.934	***
m_{t-1}	0.215		-0.038		0.041	
p_{t-1}	-0.007		0.431	***	-0.150	
y_{t-1}	-0.043		0.004		0.056	
m_{t-2}	0.033		0.025		0.007	
p_{t-2}	0.001		0.048		-0.016	
y_{t-2}	-0.069		0.021		0.028	
m_{t-3}	0.370	***	0.000		0.002	
p_{t-3}	-0.093		0.500	***	-0.063	
y_{t-3}	0.000		0.003		0.034	
c	0.008	**	0.000		-0.006	*

Notes: The asterisk marks significance at the ten (*), five (**), and one (***) percent level.

We see three key results: (i) while prices are strongly affected by excess liquidity, excess liquidity is partly absorbed by changes in money. This is most likely due to the central bank's attempts to reduce inflationary pressure; (ii) contrary to conventional wisdom, excess liquidity does not significantly impact the business cycle; and (iii) the output gap strongly affects output, which has a tendency to return to its equilibrium. Neither money supply nor prices are affected by the output gap. While all individual

coefficients that might represent a Phillips-effect are insignificant, output gap and lagged growth rates all bear the right sign. To summarize, there is weak evidence for a joint impact of the business cycle on inflation. Both, prices and money are found to be mostly driven by their own momentum, i.e. by AR processes. This corresponds to the standard findings in the literature.

Impulse response functions   A shock on demand has a contemporaneous positive impact on prices and a negative impact on money; see Figure 2.[9]

Figure 2: The economies response to a positive demand shock
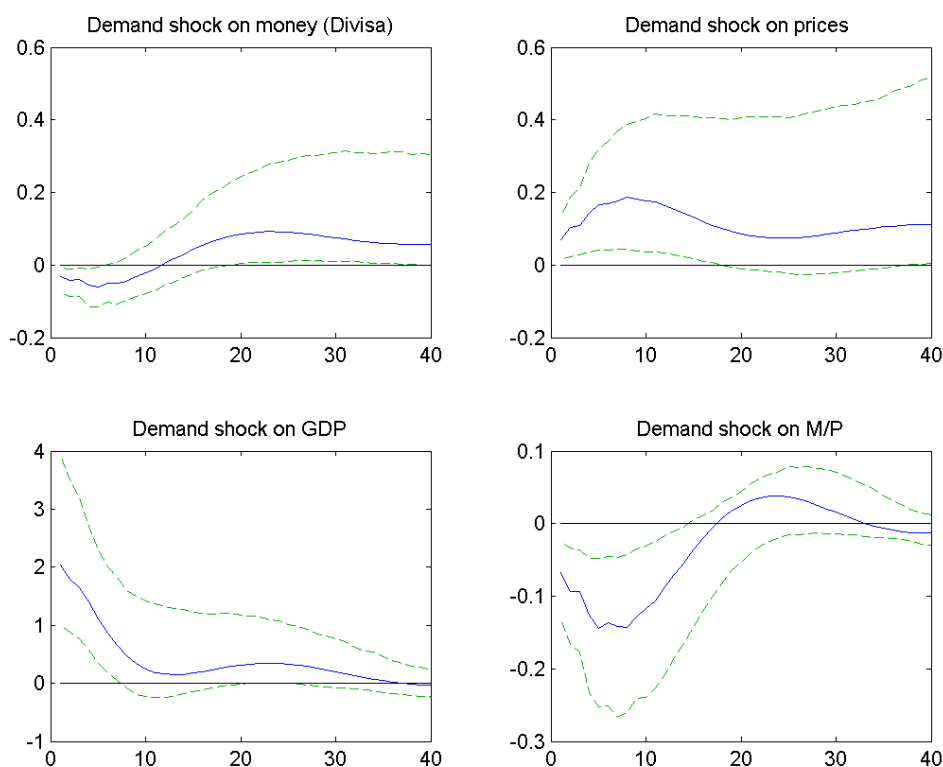
After the initial impact, the "momentum" of inflation takes over, i.e. the strong autoregressive process that is part of the inflation dynamics fosters a further increase of prices. This might be explained by adaptive expectations and the subsequent

---

[9] All impulse response functions report the consequences of a structural shock with the magnitude of one standard deviation of this specific shock, scaled with the standard deviation of residuals of the corresponding equation.

pressure on wages and commodity prices. The weakly significant consequences of a demand shock for money and prices add up to a highly significant negative impact of a demand shock on real money that lasts for several years. While the monetary conditions remain restrictive for some quarters, it is mostly money catching up to prices in this scenario and not vice versa. The momentum of inflation is strong enough to carry the price movements until monetary policy has switched to an expansionary course of action. What remains from the initial demand shock after GDP has returned to its potential, are higher prices and a correspondingly higher level of money. However, this impact is only weakly significant.

Figure 3: Response to a positive supply shock
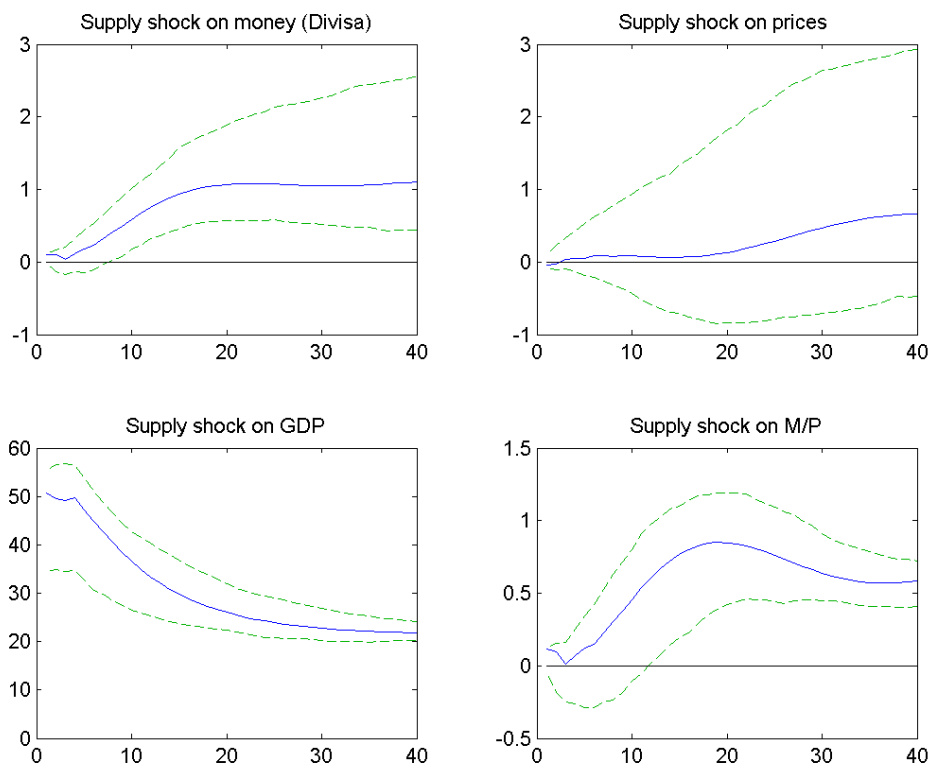
A shock to supply (see Figure 3) does not have a significant contemporaneous affect on any variable except GDP. However, the initial impact of a supply shock to GDP is larger than the supply shock itself. Hence, GDP starts declining after the initial shock. Money supply quickly starts to catch up, i.e. the central bank accommodates the increase in money demand due to higher GDP. The data suggests a slight

overreaction by the central bank leading to mild, albeit insignificant, pressure on prices. In the end, GDP remains at the new equilibrium level and real money supply is correspondingly higher as compared to the un-shocked baseline scenario.[10]

Figure 4: Response to a positive money demand shock
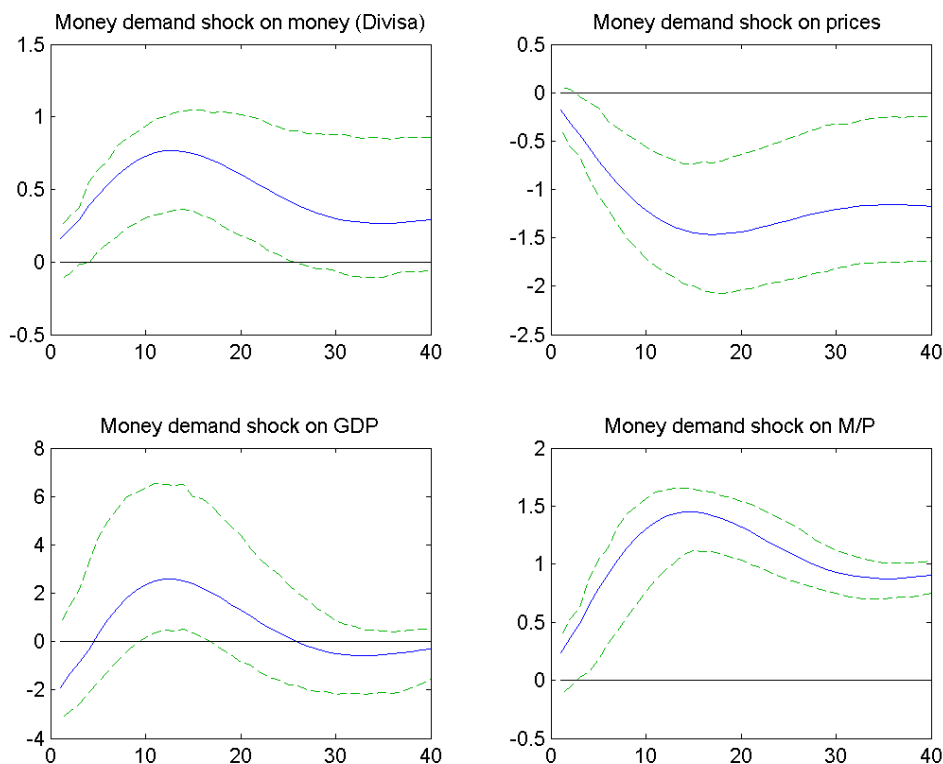
The shock to money demand, i.e. the equilibrium velocity, initially has no significant impact; see Figure 4. However, the weak reactions of GDP, prices and money in the first period quickly grow. In the third quarter after the initial shock, money supply is significantly above the baseline level, while prices are significantly higher. While the reduction in GDP can be particularly strong, even in the first period, the impact on GDP is not significant before the ninth period where we see a significantly positive

---

[10] The response of a shock to supply (i.e. potential output) seems quite large when scaled in standard deviations of the error to GDP growth. This does not indicate that the volatility of the persistent component exceeds the volatility of the cycle. However, the volatility of the GDP cycle is mostly driven by shocks to money and prices, resulting in rather small residuals in the GDP equation.

impact.[11]However, the upswing is mostly driven by the momentum of GDP after it recovers from the initial downswing. If the initial downswing is small or does not occur, the corresponding upswing is weak at best. While monetary policy quickly starts accommodating money demand, it does not compensate for the decline in price levels that occurs during the first periods after the shock. Thus, prices remain on a lower level than in the baseline scenario without shocks.

Figure 5: Response to a positive money supply shock
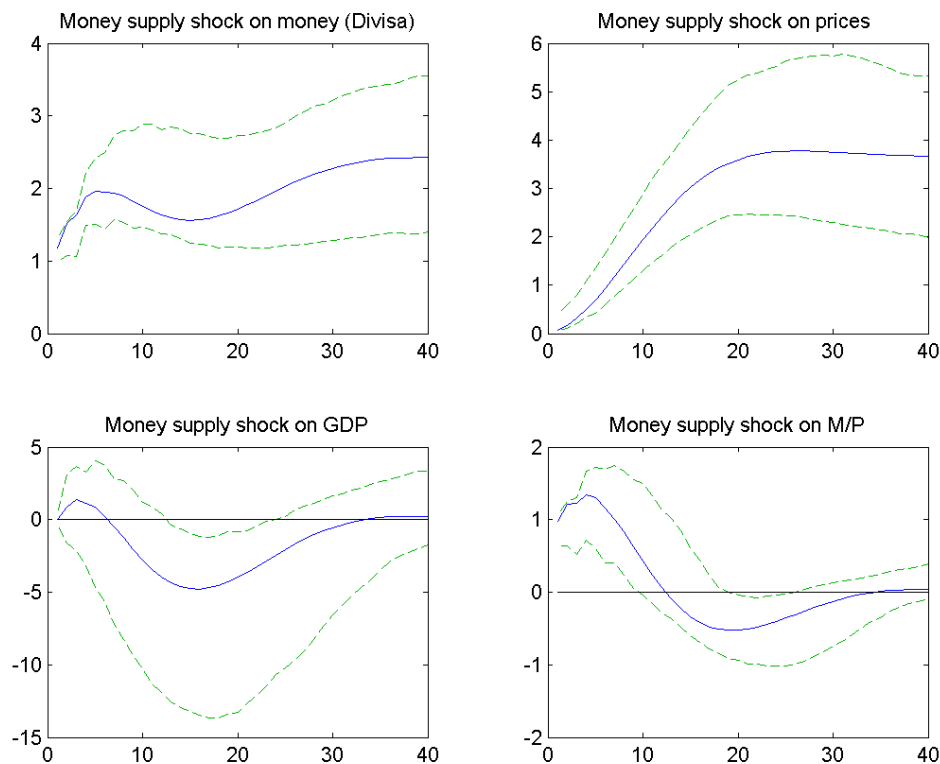
Money supply has a significantly positive contemporaneous effect on prices, as can be seen in figure 5. The initial positive impact on GDP is not significant.The following downswing, however, is significant. This downswing is only partly induced by the cyclical movement that is initiated by the shock. After money supply has risen,

---

[11] This is a puzzling result as increased money demand should lead to higher interest rates, which in turn should push down output. GDP does respond negatively to a positive money demand shock, but it is insignificant. Moreover, the response of GDP is positive after a few quarters and the response is significant. This result can be explained by change in money supply, which appears to be more important for GDP than excess liquidity.

prices quickly begin to adapt. However, due to the inertia of inflation, prices continue to rise past the level that matches the new level of money supply. This creates a lack of money supply that fosters the decline in real terms. This lack of real money becomes strongly visible when looking at the impact of money supply on real money. While the initial impact is significantly positive and remains so for 9 quarters, real money falls significantly below baseline level, once prices have started to catch up in the 14th quarter.

Figure 6: Response to a positive price shock
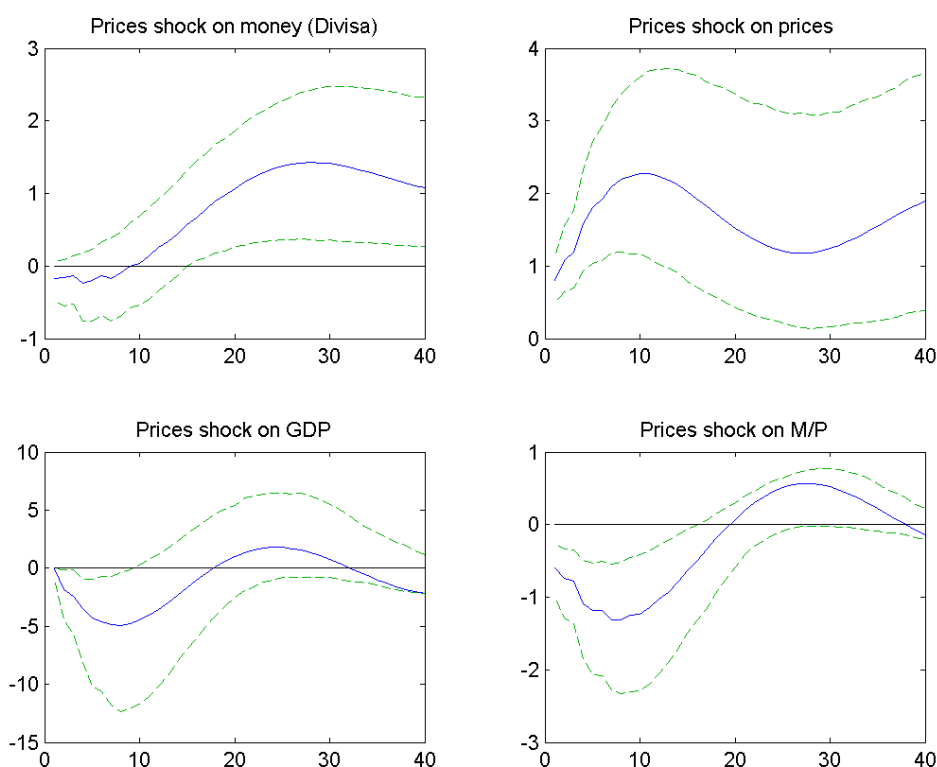
Finally, prices have a negative impact on GDP that remains significant for about two years (see Figure 6). This can only be partly attributed to the initial change in prices itself. Rather, GDP movement is driven by the ongoing price changes that follow the initial impulse due to the inertia of inflation. While there is weak evidence that the price change is mistaken for monetarily caused inflation, i.e. that the central bank reduces money supply in the face of exogenous price shocks, the price increment and the resulting increase in money demand is accommodated by money supply.

However, since money supply needs quite a while to catch up, prices have already passed their turning point and are falling again before real money has returned to its equilibrium. Thus, the remaining significantly positive impact on the price level is substantially lower than the price peak that can be observed. In the end, both prices and money remain at a higher level than the baseline model.

7. Conclusions

In this paper, we contribute to the topic of identifying monetary policy shocks in the United States economy. We propose a new unified identification scheme that identifies monetary policy, demand and supply shocks, and tracks their propagation through the economy from using a parsimonious model specification. That specification analyzes the quantity equation using Divisia money (M3), prices and GDP.

We suggest using a state space framework to disentangle velocity and output from their persistent components, equilibrium velocity and potential GDP, and to account for contemporaneous impacts of shocks to these latent variables. Then we apply a sign restriction identification scheme to the reduced form VAR given by the signal equations of the state space model. We find that a positive shock to money supply has a significant positive effect on prices, which leads us to the conclusion that the quantity theory is still relevant for the United States. Furthermore, we conclude that monetary policy in the United States did not diminish business cycles, but instead substantially contributed to fluctuations of both prices and GDP.

References

Barnett, W. and M. Chauvet (2011): How better monetary statistics could have signaled the financial crisis. *Journal of Econometrics* 161.1, pp. 6–23.

Barnett, W. A. (1980): An Application of Index Number and Aggregation Theory. *Journal of Econometrics* 14, pp. 11–48.

Barnett, W. A., U. Chae, and J. W. Keating (2005): The Discounted Economic Stock of Money with VAR Forecasting. *Annals of Finance* 2.2, pp. 229–258.

Barnett, W. A., J. W. Keating, and L. J. Kelly (2008): Toward a bias corrected currency equivalent index. *Economics Letters* 100.3, pp. 448–451.

Barnett, W. A., E. K. Offenbacher, and P. A. Spindt (1984): The New Divisia Monetary Aggregates. *The Journal of Political Economy* 92.6, pp. 1049–1085.

Barnett, W. A. and A. Serletis (2000): The Theory of Monetary Aggregation. Contributions to Economic Analysis, vol. 245. Amsterdam: Elsevier.

Bernanke, B. S. and I. Mihov (1998): Measuring Monetary Policy. *The Quarterly Journal of Economics* 113.3, pp. 869–902.

Canova, F. and G. De Nicoló (2002): Monetary disturbances matter for business fluctuations in the G-7. *Journal of Monetary Economics* 49.6, pp. 1131–1159.

Christiano, L. J., M. Eichenbaum, and C. Evans (1996): The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds. *The Review of Economics and Statistics* 78.1, pp. 16–34.

Christiano, L., M. Eichenbaum, and C. L. Evans (1999): "Monetary Policy Shocks: What have we Learned and to what End?". Handbook of Macroeconomics. Vol. 1. 1. Elsevier.

Davidson, R. and J. G. MacKinnon (2006): "Bootstrap Methods in Econometrics." Palgrave Handbooks of Econometrics. Ed. by K. Patterson and T. C. Mills. Vol. 1. Econometric Theory. Palgrave Macmillan. Chap. 25.

Dreger, C. and J. Wolters (2009): Money Velocity and Asset Prices in the Euro Area. *Empirica* 36.1, pp. 51–63.

El-Shagi, M. (2010): Did the Crisis Affect Potential Output?: *Applied Economics Letters*. forthcoming.

El-Shagi, M. and S. Giesen (2010): Money and Inflation: The Role of Persistent Velocity Movements. Discussion Paper 2/2010. Halle Institute for Economic Research.

Faust, J. (1998): The robustness of identified VAR conclusions about money. *Carnegie-Rochester Conference Series on Public Policy* 49, pp. 207–244.

Fisher, R. A. (1932): Statistical Methods for Research Workers. 4th. Oliver & Boyd.

Fry, R. and A. Pagan (2007): Some Issues in Using Sign Restrictions for Identifying Structural VARs. Working Paper 14. NCER.

Fry, R. and A. Pagan (2010): Sign Restrictions in Structural Vector Autoregressions: A Critical Review. Working Paper 22/2010. CAMA.

Gerlach, S. and L. E. O. Svensson (2003): Money and Inflation in the Euro Area: A Case for Monetary Indicators?: *Journal of Monetary Economics* 50.8, pp. 1649–1672.

Gordon, D. B. and E. M. Leeper (1994): The Dynamic Impacts of Monetary Policy: An Exercise in Tentative Identification. *The Journal of Political Economy* 102.6, pp. 1228–1247.

Hallman, J. J., R. D. Porter, and D. H. Small (1991): Is the Price Level Tied to the M2 Monetary Aggregate in the Long Run?: *American Economic Review* 81.4, pp. 841–858.

Herwartz, H. and H.-E. Reimers (2006): Long-Run Links among Money, Prices and Output: Worldwide Evidence. *German Economic Review* 7.1, pp. 65–86.

Kaufmann, S. and P. Kugler (2008): Does Money Matter for Inflation in the Euro Area?: *Contemporary Economic Policy* 26.4, pp. 590–606.

Kelly, L. J. (2009): The stock of money and why you should care. Ed. by T. B. Fomby and R. C. Hill. Emerald Group Publishing Limited.

Kelly, L. J., W. A. Barnett, and J. W. Keating (2011): Rethinking the liquidity puzzle: Application of a new measure of the economic money stock. *Journal of Banking and Finance.*

Leeper, E. and D. Gordon (1992): In search of the liquidity effect. *Journal of Monetary Economics* 29.3, pp. 341–369.

Leeper, E. M., C. A. Sims, and T. Zha (1996): What Does Monetary Policy Do?: *Brookings Papers on Economic Activity* 1996.2, pp. 1–78.

Li, H. and G. S. Maddala (1996): Bootstrapping time series models. *Econometric Reviews* 15.2, pp. 115–158.

Maddala, G. S. and S. Wu (1999): A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. *Oxford Bulletin of Economics and Statistics* 61.S1, pp. 631–652.

McCandless, G. T. and W. E. Weber (1995): Some Monetary Facts. *Federal Reserve Bank of Minneapolis Quarterly Review* 19, pp. 2–11.

Newey, W. K. and K. D. West (1987): A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55.3, pp. 703–708.

Ng, S. and P. Perron (2001): LAG Length Selection and the Construction of Unit Root Tests with Good Size and Power. *Econometrica* 69.6, pp. 1519–1554.

Orphanides, A. and R. Porter (2000): P* Revisited: Money-Based Inflation Forecasts with a Changing Equilibrium Velocity. *Journal of Economics and Business* 52.1-2, pp. 87–100.

Rudebusch, G. D. and L. E. O. Svensson (2002): Eurosystem monetary targeting: Lessons from U.S. data. *European Economic Review* 46.3, pp. 417 –442. ISSN: 0014-2921.

Sims, C. A. (1980): Macroeconomics and Reality. *Econometrica* 48.1, pp. 1–48.

Svensson, L. E. O. (2000): Does the P* Model Provide any Rationale for Monetary Targeting?: *German Economic Review* 1.1, pp. 69–81.

Tödter, K.-H. and H. Reimers (1994): P-Star as a link between money and prices in Germany. *Review of World Economics* 130.2, pp. 273–289.

Tödter, K. H. (2002): Monetary Indicators and Policy Rules in the P-Star Model. *Bundesbank Discussion Papers* 18/02.

Teles, P. and H. Uhlig (2010): Is Quantity Theory Still Alive? Working Paper No. w16393. NBER.

Uhlig, H. (2005): What are the effects of monetary policy on output? Results from an agnostic identification procedure. *Journal of Monetary Economics* 52.2, pp. 381–419.

A. Graphical Appendix

Figure A.1: Data Series

(a) Divisia Money (M3) - Level

(b) CPI - Level

(c) GDP - Level

(d) Divisia Money (M3) - Differences

(e) CPI - Differences

(f) GDP - Differences