Implications of Endogenous Money Growth for Some Tests of Superneutrality and the Fisher Effect

John W. Keating

Abstract

Superneutrality of money and the Fisher Effect are well-known theoretical propositions. Empirical tests of long-run versions of these hypotheses have sometimes been done by estimating how a variable responds to a permanent shock to inflation. Substituting inflation for money growth in a test for superneutrality is motivated by the widely-accepted Monetarist precept that “inflation is everywhere and always a monetary phenomenon.” Use of permanent shocks to inflation and money growth for testing such hypotheses has declined, in part because permanent movements in these variables have an endogenous component and so estimates are biased. But the sign of the bias may be determined using credible qualitative assumptions about the effects of structural shocks on variables. These results are used to re-examine multi-country findings from two different structural VAR models that estimate the effects of permanent inflation shocks. One finding is rejection of superneutrality for output in favor of a long-run positive output effect from permanently higher money growth. The second is rejection of the Fisher Effect in favor of nominal rates moving less than one-for-one in the long run with inflation. Both rejections are shown to be robust to endogenous money growth bias under a wide range of plausible structural assumptions. These results for real interest rates and output provide evidence in support of structural models which give rise to a Mundell-Tobin effect.

Article History: Received: 23 July 2021 / Revised: 16 January 2022 / Accepted: 22 January 2022

Keywords: JEL Classification: C32, E5

Acknowledgements

I thank the Federal Reserve Bank of Kansas City for the opportunity to serve as Visiting Scholar while working on this paper. The views expressed herein are solely those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Kansas City or the Federal Reserve System. I thank the editors, Enrique Martínez-García and Roberto Duncan, and a referee for questions and suggestions that significantly improved this paper. Discussions with Lee Smith and Vic Valcarcel have been extremely beneficial to this research. Departments of Economics at Michigan, Michigan State, Western Michigan,
Miami (Ohio), and Kansas provided very helpful comments on some of this work when it was very preliminary. The usual disclaimer applies. Support from the General Research Fund at the University of Kansas is gratefully acknowledged.

1. Introduction

Economics and finance have produced a significant number of neutrality propositions. Long-run neutrality of money occurs if a permanent change in the level of money has no long-run effect on real variables. This proposition is predicted to occur under a wide range of theories. Short-run monetary neutrality is rather uncommon, although first-generation real business cycle models have that feature.\(^1\) This paper is interested in possible long-run real effects following a permanent change in the growth rate of money. Money is said to be long-run superneutral for output if a permanent change in the money growth rate has no long-run effect on real output. A related neutrality proposition is the long-run Fisher Effect which states that a permanent change in the inflation rate has no long-run effect on the real rate of interest. That means the nominal interest rate moves one-for-one in the long-run with permanent movements in inflation according to the standard Fisher Equation. Long-run superneutrality with respect to output and the long-run Fisher Effect are the two neutrality propositions this paper is concerned with.

A large body of work examines the theoretical justification and empirical validity of these propositions. Empirical tests are crucial because theory yields an enormous range of possibilities depending on the set of structural assumptions. In the long run output may rise, fall, or be unaffected following a permanent increase in money growth. Similarly the real rate may rise, fall, or stay the same in the long run following a permanent change in the rate of inflation. Theoretical studies by Mundell (1963) and Tobin (1965) found that a permanent increase in the money growth rate results in a higher level of real output and a lower real interest rate. This combination of responses is known as a Mundell-Tobin effect. Empirical findings of Rapach (2003) provide some support for this joint effect.\(^2\)

While tests of the Fisher Effect are naturally based on inflation, inflation has also been used in tests for superneutrality. Substitution of inflation for money growth is motivated by Friedman (1996)’s classic assertion that “Inflation is always and everywhere a monetary phenomenon in the sense that it is and can be produced only by a more rapid increase in the quantity of money than in output.” Friedman was clearly referring to long-run movements in the two variables.\(^3\) While this claim may have been controversial at first, it now seems to be a guiding macroeconomic principle.

A rationale for using inflation is that money has fallen from favor for many economists. New-Keynesian Models typically ignore monetary aggregates,\(^4\) in large degree because most central

---

\(^1\) A list of seminal papers on RBC modeling would include Kydland and Prescott (1982) and Long and Plosser (1983).

\(^2\) Theory and empirical evidence on the real effects of money growth and inflation is surveyed in Orphanides and Solow (1990), Haslag (1997), and Bullard (1999).


\(^4\) Belongia and Ireland (2014) is an exception. It addresses monetary aggregation within a New-Keynesian
banks conduct policy through interest rates. There are also strong theoretical arguments for ignoring money.\textsuperscript{5} In addition, empirical work with money has frequently produced disappointing results.\textsuperscript{6}

A preference for empirical tests based on inflation may also stem from the wide range of monetary aggregates available. Along with currency, M1 includes deposit accounts that are almost entirely used for transactions purposes. Broader measures like M2, M3 and M4 include additional assets that may be directly used to transact or are highly liquid and so easily transferable into accounts which have check-writing or other features that yield liquidity. However, these additional assets are not used merely as a medium of exchange and typically have higher expected rates of return than the deposits in M1.

In addition to different levels of aggregation, there are two fundamentally different methods used to formulate monetary aggregates. One method is known as simple sum aggregation whereby each asset in an aggregate is given the same weight. The other method constructs weighted monetary aggregates where “moneyness”\textsuperscript{7} weights are provided which may differ across assets. Barnett’s Divisia measures are examples of weighted monetary aggregates.\textsuperscript{8} To satisfy the potentially wide range of professional opinions on money, one would need to use multiple measures of money and compare results. Outcomes will almost certainly vary across aggregates. In that case, a paper may get bogged down in explaining why different aggregates produce different results. In contrast, using inflation allows one to focus on the implications of the tests for long-run superneutrality and not be distracted by money measurement debates.

At one point it was popular to test superneutrality by estimating the long-run effect a permanent inflation shock has on a variable. Fisher and Seater (1993),\textsuperscript{9} Bullard and Keating (1995), and Rapach (2003) are some examples. To perform a valid test the permanent shock to inflation must be exogenous. Certainly one can imagine policy regimes in which exogeneity is plausible. For example, permanent shocks to inflation would be exogenous if a central bank targets inflation and sometimes changes the target for reasons that have nothing to do with the state of the framework with two different types of monetary assets, one called deposits which pays interest and one called currency which does not.

\textsuperscript{5}The classic reference is Woodford (2003).
\textsuperscript{6}For example, Blinder (1999) summarizes a large body of puzzling and contradictory empirical evidence for simple sum measures of money. Based on that evidence he pronounces “the death of Monetarism.”
\textsuperscript{7}Moneyness is defined in Friedman and Schwartz (1970). The concept can be thought of as how closely an asset’s liquidity is to currency’s liquidity.
\textsuperscript{8}Barnett (1978) provides a coherent theoretical framework in which he derives precise “moneyness” weights. He shows the weight for a particular asset is inversely related to its expected rate of return. See Barnett (2011) for an extended discussion of advantages, both theoretical and empirical, of weighted monetary aggregates, particularly the Divisia measures which he invented and has continued to refine.
\textsuperscript{9}The technique of Fisher and Seater (1993) is based on estimating a long-run derivative. Boschen and Otrok (1994) and Haug and Lucas (1997) used the method to test neutrality of money. Bae and Ratti (2000) use it to test superneutrality for output in data samples for Brazil and Argentina when inflation rates were very high, and find a reverse Mundell-Tobin effect, whereby in the long-run the real interest rate rises and output falls following a permanent increase in inflation. Koustas and Serletis (1999) address the Fisher Effect and discover nominal rates respond less than one-for-one in the long-run to inflation. As Equation (12) in Fisher and Seater (1993) indicates, their method for testing superneutrality is similarly based on the questionable identification assumption that permanent shocks to inflation are exogenous.
Implications of Endogenous Money Growth for Some Tests of Superneutrality and...

Another way to motivate the exogeneity assumption would be if a monetary authority follows a k% money growth rule, as advocated by Friedman (1960), with changes in k occurring only for exogenous reasons. Variations in k may come from changes in voting members in the Federal Open Market Committee. In that case, inflation would inherit exogenous permanent changes in money growth. However, it is difficult to find examples of central banks that have actually set policy with no regard for macroeconomic conditions.

Interest in using permanent shocks to inflation to test long-run neutrality propositions seems to have waned. The assumption that all permanent changes in inflation are exogenous must have played a role in this. In fact, there are various ways in which central bank behavior would make growth rates of nominal variables endogenous to shocks from the real side of the economy.

For example, oil price shocks in the 1970s are associated with persistently high inflation. A permanent increase in a relative price should not by itself have a permanent effect on the rate of inflation for an aggregate price index. However, persistent inflation may arise if a central bank elevates the money growth rate in an attempt to counteract the recessionary impact of adverse cost-push shocks. This is the point of Ireland (2007) which argues “Federal Reserve policy has systematically translated short-run price pressures set off by supply-side shocks into more persistent movements in inflation.” The fact that many nations had their highest postwar inflation rates following the 1970s oil price shocks supports this view. While the model in Ireland (2007) does not actually allow cost-push shocks to be permanent, a unit root in the relative price of oil can’t be rejected. In that case a permanent increase in this relative price could cause output to be permanently lower as industries incur expenses re-tooling production lines or re-designing products to be more energy efficient.

Another example of endogenous policy is when the monetary authority adjusts interest rates in response to the output gap. Gap measures require an estimate of full-employment or natural...
output. These measures are noisy and can be subject to substantial revision. This was particularly relevant in the 1970s when the Fed and other central banks misperceived relatively low output as the result of insufficient aggregate demand. Eventually it became clear the situation was instead a period of insufficient supply. But while this misperception continued central banks were using stimulative policy. Consequently, inflation accompanied a permanently lower level of real output. This was particularly relevant in the 1970s when the Fed and other central banks misperceived relatively low output as the result of insufficient aggregate demand. Eventually it became clear the situation was instead a period of insufficient supply. But while this misperception continued central banks were using stimulative policy. Consequently, inflation accompanied a permanently lower level of real output. Once again, higher inflation from endogenous policy goes with lower output following an adverse aggregate supply shock.

In the mid-1990s productivity growth increased for the US, and this lasted for about a decade. Standard macro theory shows that if output gradually rises to its potential level, inflation will fall. This describes how central banks have sometimes operated, in particular during the last decade of Greenspan’s long tenure as Fed Chair. A policy of this nature is called opportunistic disinflation. In essence it mirrors earlier Fed behavior except this time supply shocks were beneficial, in contrast to the 1970s experience.

There is evidence less independent central banks tend to allow more inflation. Lacking independence a central bank may be forced into printing money to help finance fiscal deficits. If an adverse real shock causes debt to accumulate and a substantial amount of debt is monetized, higher inflation may coincide with the decline in real output. Developing economies are more likely to find themselves in this situation because their central banks have historically been less independent. Also developing nations tend to have under-developed financial systems which can sometimes make it more difficult to raise new debt in world capital markets, particularly when borrowing needs are substantial. Thus less-developed countries are even more likely to respond to adverse real shocks with inflationary monetary policies.

These examples of endogenous money growth call into question the empirical tests for superneutrality or the Fisher Effect based on permanent shocks to inflation. They suggest reverse causation might be the reason for rejections. In other words, the response of monetary policy to the real economy may explain why real variables are empirically related to permanent shocks in nominal growth rates. This paper takes this idea seriously, but then dispels the notion that endogenous monetary policy explains empirical rejection of the two hypotheses.

Endogenous money growth generally biases the permanent inflation shock’s estimate away from what an exogenous change in nominal growth does to real output or the real interest rate. But if the direction of bias can be determined, an estimate’s sign might still be informative about the sign of an underlying structural effect. Specifically the paper re-examines empirical work that rejects long-run superneutrality for output and the long-run Fisher Effect in favor of a Mundell-Tobin effect. These rejections are found to be robust to endogenous money growth under a broad range of plausible structural parameters.

---

15 This argument is carefully espoused in Orphanides (2003), for example.
16 A related point is made by Kiley (2003) who found inflation was lower when productivity growth was higher. He attributes that to monetary policy allowing nominal income growth to respond less than one-for-one to changes in productivity growth.
17 Orphanides and Wilcox (2002) discusses how the approach works and when it might be useful. Meyer (2004) describes how opportunistic disinflation took place while he was a member of the Board of Governors.
In the next section important relationships between a statistical model and the structure are derived. The analysis begins with a very general economic system and an arbitrary number of endogenous variables. But to become operational, the approach requires a particular statistical model. Long-run recursive models are the focus since they have been used to address the hypotheses of interest in this paper. The relationships between the statistical model and structure are shown by mapping each of them into the unique reduced form, which in this case is a vector autoregression (VAR). Section 3 illustrates the results with a particular bivariate example. The model identifies permanent inflation shocks and that section reviews how structural assumptions were used in Keating et al. (2021) to address empirical evidence that has some bearing on superneutrality with respect to output. That paper motivates some of this work here. Section 4 examines additional evidence on the long-run SVAR coefficients obtained by Bullard and Keating (1995). This analysis provides further support for the hypothesis that superneutrality is rejected for low inflation countries in favor of output permanently rising in response to higher inflation caused by more rapid money growth. In this paper low inflation means an average in-sample inflation rate less than 11% and high inflation means 15% or more. I also entertain the hypothesis that money growth becomes more endogenous for high inflation countries and show that this does not seem a likely explanation for the differences Bullard and Keating (1995) observed between low and high inflation countries. Section 5 considers unusual features of the impulse response of output to permanent inflation shocks. Bullard and Keating (1995) find the response is nearly always lower in a high inflation country compared to a country with low inflation. Endogeneity of money growth that increases with inflation could explain this, but so could a Phillips Curve that steepens with the average rate of inflation. Another result is that countries with very high rates of inflation tend to have output responses that are never positive. This is shown to imply endogenous money growth. Overall, the evidence is consistent with the view that money growth may be more endogenous in high inflation countries. Section 6 studies the trivariate model of Rapach (2003) that was designed to test both for superneutrality and the Fisher Effect using permanent shocks to inflation. He concludes the effect of inflation on a nominal rate is less than one-for-one and that output rises for some countries and so provides evidence of Mundell-Tobin effects. My analysis will show these conclusions are actually strengthened if money growth is

---

Keating et al. (2021) and this paper build on and extend methods of Keating (2013a, b), which may be first to use credible qualitative assumptions about structure to interpret a possibly misspecified structural VAR model. Keating (2013a, b) examined long-run output neutrality of aggregate demand based on results from Blanchard and Quah (1989) decompositions of output into permanent and transitory shocks. This was done using qualitative assumptions about the effects of aggregate demand or supply on real output or the price level. No assumption on the long-run output effect from aggregate demand was made, which allowed results from the statistical model to speak on that issue.

URIbe (2022) is an interesting theoretical and empirical study in neo-Fisherian economics. In that paper’s econometric model a permanent monetary shock (one of 4 disturbances) is assumed to raise the rate of inflation and the nominal interest rate by exactly the same amount in the long run. In the short run, this shock causes the nominal interest rate, inflation, and output to rise and the real rate to fall, supporting the neo-Fisherian hypothesis that raising interest rates may stimulate the economy. Unfortunately, assumptions that a permanent monetary shock has identical long-run effects on inflation and the nominal rate and that this shock has no permanent effect on output appear at odds with the findings I report in this paper.
endogenous.  

2. The General Approach

Two representations of data are considered. One is the structural model and the other a statistical model. These representations are potentially different. Mapping both into the VAR provides a way to appreciate how the two are related since the VAR representation is unique given first and second moments of the data. This mapping will be useful whether or not the statistical model is misspecified.

A dynamic linear economic structure can be written as:

\[
\Delta X_t = \theta(L)\varepsilon_t ,
\]

where \(\theta(L) = \theta_0 + \theta_1 L + \theta_2 L^2 + \ldots = \sum_{j=0}^{\infty} \theta_j L^j\) is the structural moving average representation and \(\varepsilon_t\) is the vector of structural shocks. The number of structural shocks is assumed the same as the number of variables. Shocks are assumed to be independent with variances normalized to one: \(E \varepsilon_t \varepsilon_t' = I\). This normalization has no effect on this paper’s results, but permits a convenient way of expressing them. The variables \(X_t\) are taken as difference stationary, which is a necessary condition for there to be permanent shocks in a linear time series model with constant parameters. By recursive substitution on the structure one can determine the dynamic response of the levels of variables to structural shocks:

\[
\frac{\partial X_t}{\partial \varepsilon_{t-k}} = \sum_{j=0}^{k} \theta_j = \Phi_k ,
\]

where \(\Phi_k\) is defined as the response of \(X\) to \(\varepsilon\) shocks occurring \(k\) periods earlier. Assumptions about these effects may be used to analyze the structural content of finite horizon impulse responses from the statistical model.

Taking recursive substitution of Equation (1) to the limit yields:

\[
\lim_{k \to \infty} \frac{\partial X_t}{\partial \varepsilon_{t-k}} = \sum_{j=0}^{\infty} \theta_j = \theta(1) ,
\]

where \(\theta(1)\) evaluates \(\theta(L)\) at \(L = 1\), which indicates the long-run responses of variables to structural shocks. Ideally, one would like a statistical model to obtain precise estimates of the dynamic and long-run effects of structural shocks in Equations (2) and (3), respectively.

21Beyer and Farmer (2007) has a similar concern about addressing superneutrality when monetary policy is endogenous. They examine data on inflation, interest rates and the unemployment rate and find a single cointegrating relationship. Similar to the current paper, they map structural parameters into reduced form coefficients. After their examination of the estimates in light of theory they conclude the long-run Phillips Curve slopes upward. This implies a reverse Mundell-Tobin effect, although they don’t include output in their model. In contrast, in the context of my analysis a positive long-run relationship between inflation and the unemployment rate could be explained by the endogenous response of money growth to real shocks. For a perspective on the Phillip Curve more in line with this idea see Akerlof et al. (2000).

22It is not difficult to modify the analysis for models in which some variables are stationary without differencing.

23To estimates a structure may require stronger assumptions than the statistical model employs.
A vector autoregression (VAR) is a reduced-form representation of data. Assume a finite VAR representation of the data exists:

\[ \beta(L) \Delta X_t = e_t , \]  

with VAR coefficients given as: \( \beta(L) = I - \beta_1 L - \ldots - \beta_\ell L^\ell \) and \( \ell \) denotes the number of lags in the VAR. The vector of residuals, \( e_t \), has the same dimension as \( X_t \).

While a wide array of different assumptions have been used to identify VARs, this paper will deal only with long-run recursive statistical models. The Blanchard and Quah (1989) decomposition is one such example. Let a long-run recursive model’s moving average representation be written as:

\[ \Delta X_t = R(L)u_t , \]  

where \( R(L) = R_0 + R_1 L + R_2 L^2 + \ldots = \sum_{j=0}^{\infty} R_j L^j \) indicates the dynamic effects of the shocks, \( u_t \) is the vector of shocks, and it has the same dimension as \( X_t \). Recursive substitution on (5) for finite \( j \) yields:

\[ \frac{\partial X_t}{\partial u_{t-j}} = \sum_{i=0}^{j} R_i , \]  

which characterizes the dynamic responses of variables in levels to the model’s shocks.

The long-run impact of these shocks on variables is found by letting \( j \) go to infinity in the previous equation:

\[ \lim_{j \to \infty} \frac{\partial X_t}{\partial u_{t-j}} = \sum_{i=0}^{\infty} R_i = R(1) . \]  

\( R(1) \) is the sum of coefficients in \( R(L) \), and this represents the cumulative effect of \( u \) on \( X \). In general, \( R(1) \) may not be identical to \( \theta(1) \).

The statistical model is recursive and so is identified by assuming \( R(1) \) is a lower triangular matrix with shocks that are uncorrelated. Once again variances are for convenience normalized to 1: \( E u_t u_t' = I \). Each shock is also assumed uncorrelated with itself and other shocks at all lags and leads: \( E u_t u_{\tau} \neq 0 \forall t \neq \tau \). In essence this means the VAR has enough lags to make the residuals serially uncorrelated.

Mapping the statistical model into the VAR is done by multiplying Equation (5) by \( R_0 R(L)^{-1} \) while mapping the structure into the VAR is done by multiplying Equation (1) by \( \theta_0 \theta(L)^{-1} \). These operations require both \( R(L) \) and \( \theta(L) \) to be invertible lag polynomials, and under those conditions two useful relationships are obtained. The first one relates the VAR’s residuals to shocks, where the vector of statistical model shocks is weighted by \( R_0 \) and the vector of structural shocks is weighted by \( \theta_0 \):

\[ e_t = R_0 u_t = \theta_0 \varepsilon_t . \]  

The second useful relationship associates the VAR’s coefficients with the statistical model’s coefficients and also with structural parameters:

\[ \beta(L) = R_0 R(L)^{-1} = \theta_0 \theta(L)^{-1} . \]  

---

24 A dynamic stochastic general equilibrium (DSGE) model generally yields a vector autoregressive moving average (VARMA) representation for observable variables. Often a finite VAR representation provides a poor approximation to a VARMA. But Morris (2016) and Martinez-Garcia (2020) develop conditions under which a finite VAR representation is appropriate.
Using Equation (8), the covariance matrix of residuals is related to short-run coefficients in the statistical model and short-run parameters in the structure:

$$\Sigma_e = R_0 R_0' = \theta_0 \theta_0'.$$  \hspace{1cm} (10)

The sum of VAR coefficients matrix, $\beta(1)$, is related to the statistical model and the structure using (9):

$$\beta(1) = R_0 R(1)^{-1} = \theta_0 \theta(1)^{-1}.$$  \hspace{1cm} (11)

Identities from Equations (10) and (11) combine to yield:

$$\beta(1)^{-1} \Sigma_e \beta(1)^{-1} = R(1) R(1)' = \theta(1) \theta(1)'.$$  \hspace{1cm} (12)

The first equality indicates how $R(1)$ parameters are obtained using the appropriate Cholesky decomposition of the symmetric matrix on the left in Equation (12).

The second equality in Equation (12) shows how $R(1)$ is related to $\theta(1)$. If both are lower triangular matrices the long-run recursive model identifies the effects of all structural shocks. But this equation holds even when the structure is not recursive, and in that case it provides the mapping between structural parameters and coefficients in the long-run recursive model. These relationships along with information about long-run effects of structural shocks will be used to derive useful information about the underlying structure.

One may also be interested in the relationship between the dynamic responses from the statistical model and the structural responses. From the second equation in (9) obtain:

$$R(L) = \theta(L) \theta(1)^{-1} R(1).$$  \hspace{1cm} (13)

or equivalently:

$$R_i = \theta_i \theta(1)^{-1} R(1) \quad \text{for} \quad i = 0, 1, 2, ...$$  \hspace{1cm} (14)

Combining the last expression with Equation (6) and recalling the definition of structural effects in Equation (2) yields the relationship between statistical model impulse responses and structural impulse responses:

$$\frac{\partial X_t}{\partial u_{t-j}} = \sum_{i=0}^j R_i = \sum_{i=0}^j \theta_i \theta(1)^{-1} R(1) = \Phi_j \theta(1)^{-1} R(1).$$  \hspace{1cm} (15)

If $R(1) = \theta(1)$, the statistical model will identify the structure’s impulse responses. Otherwise, the statistical model’s responses are linear combinations of structural responses with weights given by $\theta(1)^{-1} R(1)$. If $\theta(1)$ is lower triangular, these parameters are related to $R(1)$ from the second identity in Equation (12). These relationships are essential for understanding how the long-run recursive statistical model’s coefficients are related to the structure. They are also instrumental in using Equation (15) to determine how the model’s impulse responses are related to structural responses.

---

25 This is because a recursive model is unique. Hamilton (1994, p. 91) provides a proof of this claim.

26 There exists a class of structures that permit the long-run recursive model to identify responses for a subset of structural shocks. For more details, see Keating (2002).
3. A Bivariate Example

A bivariate model provides the simplest example of how one can use the relationships derived in Section (2). Decompositions of inflation into permanent and transitory shocks are relevant to the main concerns of this paper. The choice of second variable depends on the question at hand. Following Bullard and Keating (1995) that variable is real output. Therefore, $X_t = (\pi_t, y_t)'$, and the statistical model can be written as:

$$
\begin{bmatrix}
\Delta \pi_t \\
\Delta y_t
\end{bmatrix} =
\begin{bmatrix}
R^{\pi P}(L) & R^{\pi T}(L) \\
R^{y P}(L) & R^{y T}(L)
\end{bmatrix}
\begin{bmatrix}
u_t^P \\
u_t^T
\end{bmatrix}.

(16)
$$

The temporary inflation shock’s long-run effect on $\pi$ is zero by construction:

$$R^{\pi T}_t = \sum_{j=0}^{\infty} R^{\pi T}_{j} = 0 .$$

The first shock in the model is allowed to have a permanent effect on inflation, the second one does not do that, and both shocks may affect output in the long run. Thus, $R(1)$ is a lower triangular matrix:

$$R(1) =
\begin{bmatrix}
R^{\pi P} & 0 \\
R^{y P} & R^{y T}
\end{bmatrix}.

(17)
$$

This is a Blanchard and Quah (1989) decomposition, but one that identifies the responses of inflation and output to permanent and transitory shocks to inflation rather than output.

Now consider a structure:

$$\begin{bmatrix}
\Delta \pi_t \\
\Delta y_t
\end{bmatrix} =
\begin{bmatrix}
\theta^{\pi \mu}(L) & \theta^{\pi \lambda}(L) \\
\theta^{y \mu}(L) & \theta^{y \lambda}(L)
\end{bmatrix}
\begin{bmatrix}
\mu_t \\
\lambda_t
\end{bmatrix},

(18)
$$

where the two variables are driven by exogenous shocks to money growth and aggregate supply, $\varepsilon_t = (\mu_t, \lambda_t)'$, and lag polynomials $\theta^{v s}(L)$ for $v = \pi, y$ and $s = \mu, \lambda$ represent dynamic responses of variables to structural disturbances. The dynamic responses of variables to shocks comes from the bivariate version of (2)

$$\frac{\partial X_t}{\partial \varepsilon_{t-j}} =
\begin{bmatrix}
\Phi^{\pi \mu}_j & \Phi^{\pi \lambda}_j \\
\Phi^{y \mu}_j & \Phi^{y \lambda}_j
\end{bmatrix},

(19)
$$

where

$$\Phi^{v s}_j = \sum_{i=0}^{j} \theta^{v s}_{i} , \text{ for } v = \pi, y \text{ and } s = \mu, \lambda .$$

The long-run parameter matrix is specified as:

$$\theta(1) =
\begin{bmatrix}
\alpha_{\pi \mu} & \alpha_{\pi \lambda} \\
\alpha_{y \mu} & \alpha_{y \lambda}
\end{bmatrix}.

(20)
$$

If aggregate supply has no long-run effect on inflation, then \( \alpha_{\pi\lambda} = 0 \), \( \theta(1) \) is lower triangular, and thus the statistical model and the structure are equivalent. But more generally, that restriction on the structure may not hold, and then each coefficient from the statistical model is a function of structural parameters. Insert \( R(1) \) from (17) and \( \theta(1) \) from (20) into the second identity from Equation (12), and solve for the statistical model’s long-run coefficients:

\[
R_{\pi P} = \sqrt{\alpha_{\pi\mu}^2 + \alpha_{\pi\lambda}^2} \quad R_{yP} = \frac{\alpha_{\pi\mu}\alpha_{y\mu} + \alpha_{\pi\lambda}\alpha_{y\lambda}}{\sqrt{\alpha_{\pi\mu}^2 + \alpha_{\pi\lambda}^2}} \quad R_{yT} = \frac{\alpha_{\pi\mu}\alpha_{y\lambda} - \alpha_{\pi\lambda}\alpha_{y\mu}}{\sqrt{\alpha_{\pi\mu}^2 + \alpha_{\pi\lambda}^2}}.
\]  

(21)

These relationships are essential for understanding the structural implications of a potentially misspecified statistical model of permanent inflation shocks.\(^{28}\)

The long-run percentage change in output associated with a 1 percentage point increase in the inflation rate is obtained from \( \frac{R_{yP}}{R_{\pi P}} \), and this is related to structural parameters using Equation (21):

\[
\frac{R_{yP}}{R_{\pi P}} = \frac{\alpha_{y\mu} + \alpha_{\pi\lambda}\alpha_{y\lambda}}{\alpha_{\pi\mu}^2 + \alpha_{\pi\lambda}^2}.
\]  

(22)

Dividing numerator and denominator by \( \alpha_{\pi\mu}^2 \) yields:

\[
\frac{R_{yP}}{R_{\pi P}} = \frac{\frac{\alpha_{y\mu}}{\alpha_{\pi\mu}} + \frac{\alpha_{\pi\lambda}\alpha_{y\lambda}}{\alpha_{\pi\mu}^2}}{1 + \frac{\alpha_{\pi\lambda}^2}{\alpha_{\pi\mu}^2}}.
\]  

(23)

Rewriting \( \frac{R_{yP}}{R_{\pi P}} \) in this way isolates \( \frac{\alpha_{y\mu}}{\alpha_{\pi\mu}} \). This ratio of structural parameters represents the long-run output effect of a one percentage point exogenous increase in money growth, which is what one would like to obtain with the long-run recursive model. If money growth is not endogenous (\( \alpha_{\pi\lambda} = 0 \)), permanent inflation shocks can be used to test for supernutrality.

But if that key identification assumption is false, Equation (23) indicates there are two terms that may bias the estimate from what economists are trying to obtain. To interpret this equation requires additional information about the structure. Economic reasoning and theory often provide the sign, if not the precise magnitude, of a structural effect. These inequality constraints can be used to determine the structural implications of the effects of permanent inflation shocks. Sign restrictions have been widely used to identify VARs.\(^{29}\) However, the approach taken here is fundamentally different. Sign restrictions are not used for estimation purposes. They are instead used in an attempt to provide structural interpretations of coefficients and impulse responses from a misspecified long-run recursive statistical model.

\(^{28}\)Notice positive square roots have been taken. This is most often the option chosen in applied work. For example, in the bivariate case at hand, choosing the positive square root in both cases implies the model’s impulses will have one shock that permanently raises inflation and may also affect output in the long run (but doesn’t have to) and another shock that raises output in the long run but has no long-run effect on inflation. For the results in this paper, the choice of a positive or a negative square root is completely irrelevant.

Three intuitive inequalities on long-run structural parameters are assumed to hold:

S1: $\alpha_{\pi \mu} > 0$
S2: $\alpha_{y \lambda} > 0$
S3: $\alpha_{\pi \lambda} < 0$.

S1 asserts that a permanent exogenous increase in the growth rate of money permanently raises the inflation rate. I am unaware of any theory in which this inequality would be rejected. Assumption S2 states that an exogenous permanent increase in aggregate supply will raise output in the long run. This assumption is at least as incontrovertible as S1. Assumption S3 states that a permanent beneficial aggregate supply shock will result in a permanently lower rate of inflation while an adverse supply shock causes inflation to rise permanently. Arguments for endogenous money growth made in the Introduction support this assumption.

Keating et al. (2021) used this set of assumptions to interpret results from the bivariate model of permanent inflation shocks. They show downward bias arises for all positive values of this long-run effect. Bias also occurs for a range of negative values and if the parameter is zero. Hence, negative estimates of $\frac{R_{y\pi}}{R_{e\pi}}$ are uninformative, as they may occur if the structural response of output is positive, zero, or negative. On the other hand positive estimates of $\frac{R_{y\pi}}{R_{e\pi}}$ are informative. Downward bias from endogenous money growth means $\frac{\alpha_{y\pi}}{\alpha_{\pi\mu}}$ is expected to be even larger than the positive estimate. Bullard and Keating (1995) obtained positive estimates for 9 of 10 countries with average inflation rates under 11%. In 4 of the countries, the effect is statistically significant and in 3 others the effect is very nearly significant. Using the aforementioned theoretical findings, Keating et al. (2021) interpret this empirical evidence as support for the hypothesis that a permanent increase in money growth can raise output in the long run for a group of low inflation countries.

The United States is the only low inflation country for which Bullard and Keating (1995) estimate a negative output effect. This motivated Keating et al. (2021) to develop a trivariate model that controls for endogenous money growth when estimating the long-run output effect from permanent inflation. They find their control yields a significant positive estimate. That means all of the low inflation countries in Bullard and Keating (1995) provide some evidence of a long-run positive output effect from permanently higher inflation.


Bullard and Keating (1995) obtained a variety of other empirical findings, some of which were given structural explanations. However, endogenous money growth raises questions about the legitimacy of those explanations. This section re-examines that evidence, extending Keating et al. (2021)’s investigation of the structural implications of the bivariate SVAR model’s results.

30 While one might want to go further and assume $\alpha_{\pi\mu} = 1$ based on Friedman’s inflation principle, that would be a mistake! The unit variance normalization for each shock in the model implies that an $\alpha_{vs}$ parameter actually is equal to the structural effect of shock $s$ on variable $v$ multiplied by the standard deviation of shock $s$.

31 Keating et al. (2021) also examined potential sources of bias in their trivariate model and concluded their findings are robust to those sources.
4.1 Endogenous Money, Inflation, the Effects of Permanent Inflation Shocks

Bullard and Keating (1995) find that 9 of 10 low inflation countries obtain positive estimates while all 6 high inflation countries obtain non-positive estimates. In fact, Argentina had the highest in-sample inflation rate (143.6%) and was the only country with a significantly negative long-run output response to a permanent increase in inflation. They argue this evidence is consistent with the model of Azariadis and Smith (1996) whereby financial frictions cause low inflation economies to experience Mundell-Tobin effects which dissipate when inflation reaches a certain point. Would bias from endogenous money growth spoil that interpretation?

To address this question, take the derivative of Equation (23) with respect to $\alpha_{y\mu}$ holding fixed all other structural parameters:

$$\frac{d}{d\alpha_{y\mu}} \left( \frac{R_{yP}}{R_{\pi P}} \right) = \frac{\alpha_{\pi\mu}}{\alpha_{\pi\mu}^2 + \alpha_{\pi\lambda}^2} > 0. \tag{24}$$

This is positive since the denominator is clearly positive and by assumption S1 so is the numerator. Thus if $\alpha_{y\mu}$ varies with inflation in accord with the model of Azariadis and Smith (1996) that pattern should show up even if the estimates are biased. In fact if $\alpha_{\pi\mu}$ and $\alpha_{\pi\lambda}$ are the same for all countries this derivative will be constant across countries. In that case, a cross-country pattern in structure would be replicated in the statistical model up to a scale factor.

4.2 Can Endogeneity Increasing with Inflation Explain Cross-Country Findings?

Countries with less independent central banks are prone to higher rates of inflation. That is the lesson from evidence presented in Alesina and Summers (1993). A less independent central bank is more likely to be pressed into responding to adverse economic circumstances. This suggests the $\alpha_{\pi\lambda}$ parameter is likely more negative in countries with higher rates of inflation.

Could the cross-country pattern of estimates in Bullard and Keating (1995) be explained by this hypothesis of systemic variation in $\alpha_{\pi\lambda}$ with inflation? To address this question take the derivative of Equation (23) with respect to $\alpha_{\pi\lambda}$:

$$\frac{d}{d\alpha_{\pi\lambda}} \left( \frac{R_{yP}}{R_{\pi P}} \right) = \frac{\alpha_{\pi\mu}^2 \alpha_{y\lambda} - \alpha_{\pi\lambda}^2 \alpha_{y\lambda} - 2\alpha_{y\mu} \alpha_{\pi\mu} \alpha_{\pi\lambda}}{(\alpha_{\pi\mu}^2 + \alpha_{\pi\lambda}^2)^2}. \tag{25}$$

If this derivative is positive, then the relationship between estimates and inflation could be explained by $\alpha_{\pi\lambda}$ becoming more negative as inflation increases. However, the sign is ambiguous. The denominator is positive, but the numerator consists of terms that are positive, negative, and ambiguous, respectively. The last term in the numerator takes the sign of $\alpha_{y\mu}$. Thus variation in $\alpha_{\pi\lambda}$ appears an unlikely explanation.

And what if $\alpha_{y\mu}$ turns negative once inflation reaches a critical rate? This situation is quite likely given the harsh effects hyperinflation can have on output. With this parameter less than

---

32 Lagos and Rocheteau (2005) obtain a similar finding in a search model, although welfare falls with inflation. Their model has posted prices and directed search by buyers.
zero and $\alpha_{\pi\lambda}$ even more negative in countries with high rates of inflation this derivative shrinks and is increasingly likely to be negative. I conclude that the way these long-run estimates vary with inflation can not be explained by money growth endogeneity that rises with inflation.


Equation (15) indicates how each of the statistical model’s impulse responses is a linear combination of structural impulse responses with weights determined by $\theta(1)^{-1}R(1)$. Combine this equation with $R(1)$ from (17) along with the relationships in (21) and $\theta(1)$ from (20), and after some algebra:

$$\theta(1)^{-1}R(1) = \begin{bmatrix} \alpha_{\pi\mu} & -\alpha_{\pi\lambda} \\ \alpha_{\pi\lambda} & \alpha_{\pi\mu} \end{bmatrix} \frac{\alpha_{\pi\mu}}{\sqrt{\alpha_{\pi\mu}^2 + \alpha_{\pi\lambda}^2}}. \quad (26)$$

This equation combined with (15) and (19) determines how the model’s impulse responses are related to the structural impulse responses and the long-run structural parameters:

$$\frac{\partial X_t}{\partial u_{t-k}} = \begin{bmatrix} \Phi_{\pi\mu}^k & \Phi_{\pi\lambda}^k \\ \Phi_{\mu\mu}^k & \Phi_{\mu\lambda}^k \end{bmatrix} \begin{bmatrix} \alpha_{\pi\mu} & -\alpha_{\pi\lambda} \\ \alpha_{\pi\lambda} & \alpha_{\pi\mu} \end{bmatrix} \frac{\alpha_{\pi\mu}}{\sqrt{\alpha_{\pi\mu}^2 + \alpha_{\pi\lambda}^2}}, \quad (27)$$

from which the response of output to a permanent inflation shock is:

$$\frac{\partial y_t}{\partial u_{t-k}} = \frac{\Phi_{\mu\mu}^k \alpha_{\pi\mu} + \Phi_{\mu\lambda}^k \alpha_{\pi\lambda}}{\sqrt{\alpha_{\pi\mu}^2 + \alpha_{\pi\lambda}^2}}. \quad (28)$$

To interpret this impulse response from the statistical model assumptions S1 and S3 must be augmented by plausible assumptions about the dynamic effects of structural shocks. Assume the initial output effect from an exogenous positive money growth shock is positive:

S4: $\Phi_{\mu\mu}^k > 0$ for $0 \leq k \leq K$ for some finite $K > 0$.

Also, assume output responds positively to a beneficial aggregate supply shock:

S5: $\Phi_{\mu\lambda}^k > 0$ for $k \geq 0$.

Assumption S4 imposes no restriction on the sign or magnitude of the long-run response of output to an exogenous shock to money growth. Assumption S5 rules out cases in which output falls initially in response to a technological improvement. Assumptions about dynamic responses of inflation to structural shocks are not necessary for interpreting the response of output to a permanent inflation shock.

---

33Since Bullard and Keating (1995) estimates are done using annual data this assumption is that output responds positively within a year following an exogenous increase in money growth.

34Basu et al. (2006) argue output falls in the short-run following a positive technology shock, but that decline is relatively brief and does not persist for a year. Hence, in annual data this negative effect would not be observable.
Impulse responses from Bullard and Keating (1995) are examined in light of these structural assumptions. Most of these responses were removed from the final version of the paper, but they can be found in an on-line St. Louis Federal Reserve Bank working paper.\textsuperscript{35}

5.1 Output Responses to a Permanent Shock are Related to Inflation

Impulse responses of output to a permanent inflation shock in Bullard and Keating (1995) tend to be higher in low inflation countries. In fact, the entire reported finite horizon impulse response for a high inflation country is typically below that of a low inflation country.

Can this empirical tendency be explained by $\alpha_{\pi \lambda}$ being more negative in high inflation countries? This hypothesis is reasonable to consider since, as was noted earlier, high inflation countries on average have less independent central banks. To answer this question take the partial derivative of Equation (28) with respect to $\alpha_{\pi \lambda}$:

\[
\frac{\Phi_{y\lambda}^{\mu} - \Phi_{\mu}^{\mu} \alpha_{\pi \mu} \alpha_{\pi \lambda}}{\left(\alpha_{\pi \mu}^2 + \alpha_{\pi \lambda}^2\right)^{3/2}}.
\]

(29)

Given S1, S3, S4, and S5 this derivative is positive. That means that the dynamic response of output to a permanent shock to inflation shifts downward as $\alpha_{\pi \lambda}$ becomes more negative. Thus, variation of impulse responses between low and high inflation economies might be explained by $\alpha_{\pi \lambda}$ falling with inflation, possibly because higher inflation countries have less independent central banks. However, this hypothesis is not the only plausible explanation.

The dynamic responses of output to exogenous money growth shocks would be expected to shift down with inflation if the Phillips Curve becomes steeper with inflation. Ball et al. (1988) makes the case: “In countries with low inflation, the short-run Phillips curve is relatively flat—fluctuations in nominal aggregate demand have large effects on output. In countries with high inflation, the Phillips curve is steep—fluctuations in demand are reflected quickly in the price level.” To determine if the impulse response evidence might be explained by this hypothesis, take the derivative of (28) with respect to $\Phi_{y\mu}^{\mu}$:

\[
\frac{\alpha_{\pi \mu}}{\sqrt{\alpha_{\pi \mu}^2 + \alpha_{\pi \lambda}^2}} > 0 \text{ given S1}.
\]

(30)

This derivative is positive which proves the dynamic response of output to a permanent inflation shock moves with the response of output to an exogenous money growth shock.\textsuperscript{36} Thus, a Phillips Curve that becomes steeper with inflation provides a second plausible explanation for output responses to permanent inflation shocks being lower for high inflation countries.

\textsuperscript{35}At: https://research.stlouisfed.org/wp/more/1994-011. Note the working paper’s title is “Superneutrality in Postwar Economies.” To satisfy a referee, the published version had to remove “Superneutrality” and some of its discussion. The primary criticism was that permanent shocks to inflation may not be exogenous.

\textsuperscript{36}In the 21st Century inflation’s behavior relative to the state of the business cycle has caused a number of economists to question Phillips Curves.
5.2 Non-positive Output Responses are Common in High-Inflation Economies

The output response to a permanent inflation shock is almost never negative in the low inflation countries. The United States is the primary exception. The high inflation country responses are typically very different. For Argentina, Chile, and Costa Rica the output response to a permanent inflation shock is never positive and is nearly always negative. For Iceland the response is zero or slightly negative.37

If the response in (28) is always non-positive, that implies an inequality for the long-run effect of aggregate supply:

\[ \alpha_{\pi\lambda} \leq \frac{-\phi_k^{\mu\mu}}{\Phi_k^{\mu\lambda}} \alpha_{\pi\mu} \]  

(31)

Assumptions S1, S5, and S6 guarantee that the right side is negative. Hence these non-positive output responses to a permanent increase in inflation imply that money growth was endogenous to real shocks for high inflation countries.

Historically, cases of hyperinflation or exceptionally high inflation have always been associated with large budget deficits and a monetary authority that had little if any independence. In fact, the two exceptionally high inflation countries in the sample, Chile and Argentina, ran huge budget deficits in the 70s and 80s, and those decades covered a large portion of the sample period used by Bullard and Keating (1995). These deficits were in large part a consequence of the world-wide recessions caused by adverse oil price shocks, when clearly both countries printed a great deal of money to finance public debt. The experience of these two very high inflation countries adds further support to the result that non-positive impulse responses imply \( \alpha_{\pi\lambda} \) less than zero.

6. Rapach (2003)’s Tests of Two Neutrality Propositions

Rapach (2003) estimates a trivariate VAR model with inflation, a nominal interest rate, and real output using post-World War II data for 14 countries,38 jointly testing the hypotheses that these inflation shocks have: (1) no long-run effect on output; and (2) a one-for-one long-run effect on the nominal interest rate. The model is a multivariate Blanchard and Quah (1989) decomposition with inflation ordered first in the long-run recursive model. Rapach (2003)’s output results are similar to Bullard and Keating (1995). A permanent positive shock to inflation

---

37 The other two high-inflation countries behave differently. Portugal’s response resembles that of the US. Perhaps that is not too surprising given that it had the lowest rate for the high inflation countries. Mexico’s results are more puzzling. Initially, the Mexican response is positive. But its confidence bounds continually expand as the response horizon increases. Also, the response of inflation to a temporary inflation shock appears like it may have a complex non-stationary root. If that is true this response will never settle down to zero, in contradiction to the long-run restriction. Hence, this model has problems accounting for the relationship between inflation and output in Mexico, at least for that particular sample period.

38 Data from the period before World War II might be ignored if inflation is stationary. Barsky (1987), for example, argues inflation must be integrated of order 1 to perform a statistical test of the Fisher Effect. Plots of inflation in the period prior to World War II suggest inflation may have been stationary. Alternatively one could use time varying parameter models, such as Stock and Watson (2007) or Martínez-García (2018), which allow the variance of permanent shocks to fall in periods where inflation seems to be closer to a stationary process.
causes output to be permanently higher in most countries, some of these estimates are statistically significant, and quite a few are nearly significant. Furthermore Rapach (2003) finds in nearly all cases that the nominal interest rate moves significantly less than one-for-one with a permanent shock to inflation. He interprets these two sets of results as evidence of Mundell-Tobin effects.

Rapach (2003)’s long-run structural assumptions are: inflation is only affected by exogenous inflation shocks, the real interest rate is only affected by preference shocks ($\rho_t$), inflation shocks and preference shocks affect the nominal rate, and real output may respond to inflation shocks, preference shocks, and technology shocks. He orders variables as $X_t = \left( \pi_t, i_t, y_t \right)'$ and the structural shocks as: $\varepsilon_t = \left( \mu_t, \rho_t, \lambda_t \right)'$. Given that set of structural assumptions, the model can address long-run superneutrality with respect to output and the long-run Fisher Effect.

The statistical model has shocks given by $u_t = \left( u^P_t, u^2_t, u^3_t \right)'$, where $u^P_t$ is a permanent shock to inflation and the other two shocks are constrained to have only temporary effects on the inflation rate. Since the statistical model and the unrestricted structure are lower triangular matrices of the same form, $u_t = \varepsilon_t$ and $R(1) = \theta(1)$. Expanding this matrix expression yields:

\[
\begin{bmatrix}
R_{\pi P} & 0 & 0 \\
R_{i P} & R_{i2} & 0 \\
R_{y P} & R_{y2} & R_{y3}
\end{bmatrix} = \begin{bmatrix}
\alpha_{\pi \mu} & 0 & 0 \\
\alpha_{i \mu} & \alpha_{\pi \rho} & 0 \\
\alpha_{y \mu} & \alpha_{y \rho} & \alpha_{y \lambda}
\end{bmatrix}.
\]

Therefore:

\[
\frac{R_{y P}}{R_{\pi P}} = \frac{\alpha_{y \mu}}{\alpha_{\pi \mu}},
\]

and so testing if this ratio of parameters is equal to zero addresses the hypothesis that money is superneutral with respect to output.

Furthermore:

\[
\frac{R_{i P}}{R_{\pi P}} = \frac{\alpha_{i \mu}}{\alpha_{\pi \mu}},
\]

and so testing this ratio of parameters is equal to one is a test of the Fisher effect on interest rates.\(^{39}\)

While Rapach (2003) defines $\rho$ as a shock to preferences, this paper reinterprets it as an (adverse) exogenous shock to government savings. A major reason for altering the interpretation comes from a standard result in consumer optimization theory. If the utility function in a dynamic macro model is augmented with preference shocks, this is typically done by multiplying the discount rate by a function of that shock. When this shock has a unit root its long-run effect on the real interest rate disappears from the linearized Euler Equation.\(^{40}\) Hence, based on standard dynamic macroeconomic theory it is internally inconsistent for a permanent preference shock to have a permanent effect on the real interest rate.

An alternative model for the real interest rate can be derived from a long-run equilibrium condition that equates savings to investment. A more detailed exposition of the simple structural framework is developed in Appendix A. It features a common Cobb-Douglas production function,

\(^{39}\)Rapach (2003) did not examine the effects of preference and technology shocks, but could have if the theory was correct.

\(^{40}\)Equation (18) in Ireland (2004) provides an early example of this general result.
the demand for capital derived from that production function, and an assumption that long-run desired savings is a function of real output, the real interest rate and a shock to government savings. If income elasticities for savings and investment are equal, the savings shock is the only disturbance which has a long-run effect on the real rate. This is the intuition for reinterpreting the $\rho$ shock. Importantly, altering the source of this shock does not invalidate the identification restrictions or the results in Rapach (2003).

Also, interpreting $\rho$ as an adverse shock to government savings makes this paper’s long-run structure very similar to that of Ahmed and Rogers (2000). Each paper models three permanent structural disturbances, with money growth and technology shocks common in both. In contrast, Ahmed and Rogers (2000)’s third shock is associated with government expenditures. They find that this government spending shock plays an important role in macroeconomic fluctuations. My structural model allows $\rho$ shocks to also account for exogenous shocks to taxation.\footnote{The model of Ahmed and Rogers (2000) has cointegration and more variables than Rapach (2003). King et al. (1991) also has cointegration and 3 similar shocks. In contrast to each of these papers, I do not assume that the statistical model necessarily identifies a structure.}

The next modification to Rapach (2003) is to permit endogenous money growth. Once more inflation is allowed to endogenously respond to aggregate supply shocks in the long run. Economic reasoning also suggests a possible long-run relationship between money growth and $\rho$ shocks. For example, money growth may increase if there is a permanent decline in government savings and the central bank monetizes some portion of rising government debt.

I also consider the possibility that nominal interest rates may not respond one-for-one to inflation in the long run. To do so assume that when the rate of inflation increases by 1 percentage point, the nominal interest rate increases by $\eta$ percentage points. This neither rules out nor imposes the Fisher Effect.

These modifications to Rapach (2003) yield:

$$\theta(1) = \begin{bmatrix} \alpha_{\pi\mu} & \alpha_{\pi\rho} & \alpha_{\pi\lambda} \\ \eta \alpha_{\pi\mu} & \eta \alpha_{\pi\rho} + \alpha_{\tau\rho} & \eta \alpha_{\pi\lambda} \\ \alpha_{y\mu} & \alpha_{y\rho} & \alpha_{y\lambda} \end{bmatrix}. \quad (35)$$

The first row of $\theta(1)$ allows inflation to potentially respond in the long run to money growth, government savings, and technology shocks. The second row assumes the real rate responds to the government savings shock, and the nominal rate responds to the real rate and a permanent change in inflation, but without imposing the Fisher Effect. The third row indicates that each of the three permanent shocks potentially may affect real output in the long run.

Combining $R(1)$ from Equation (12):

$$\begin{bmatrix} R_{\pi P} & 0 & 0 \\ R_{1 P} & R_{12} & 0 \\ R_{y P} & R_{y2} & R_{y3} \end{bmatrix} \begin{bmatrix} R_{\pi P} & R_{1 P} & R_{y P} \\ R_{1 P} & R_{12} & R_{y2} \\ R_{y P} & R_{y2} & R_{y3} \end{bmatrix} = \begin{bmatrix} \alpha_{\pi\mu} & \alpha_{\pi\rho} & \alpha_{\pi\lambda} \\ \eta \alpha_{\pi\mu} & \eta \alpha_{\pi\rho} + \alpha_{\tau\rho} & \eta \alpha_{\pi\lambda} \\ \alpha_{y\mu} & \alpha_{y\rho} & \alpha_{y\lambda} \end{bmatrix} \begin{bmatrix} \alpha_{\pi\mu} & \eta \alpha_{\pi\mu} & \alpha_{y\mu} \\ \alpha_{\pi\rho} & \eta \alpha_{\pi\rho} + \alpha_{\tau\rho} & \alpha_{y\rho} \\ \alpha_{\pi\lambda} & \eta \alpha_{\pi\lambda} & \alpha_{y\lambda} \end{bmatrix}. \quad (36)$$

allows for a solution of each parameter from the long-run recursive statistical model in terms of structural parameters. As may already be apparent, each coefficient in the recursive statistical model is a function of multiple structural parameters even if the null hypotheses, $\eta = 1.0$ and...
\( \alpha_{\mu} = 0 \), are true. Hence, once again a long-run recursive statistical model is incapable of identifying structural effects.

Previous assumptions S1, S2, and S3 on the effects of technology will still be used, but now assumptions about the long-run effects of \( \rho \) shocks are also required:\(^{42}\)

\[
\begin{align*}
S6: & \quad \alpha_{\pi \rho} > 0 \\
S7: & \quad \alpha_{r \rho} > 0 \\
S8: & \quad \alpha_{y \rho} < 0 .
\end{align*}
\]

Assumption S7 states that an exogenous decline in government savings raises the real interest rate. S8 assumes a reduction in government savings reduces the capital stock causing output to fall. Assumption S6 asserts that a reduction in government savings increases its borrowing, and if the central bank monetizes debt the rate of inflation may increase.\(^{43}\)

Rapach (2003) estimates the trivariate model using annual data for 14 countries but due to data limitations can only do so with quarterly data for 5 of those countries. The following two sub-sections address the implications for the two neutrality propositions from this misspecified empirical model.

### 6.1 Rejection of the Fisher Effect is Robust to Endogenous Money

Rapach (2003)’s estimates of \( \frac{R_{iP}}{R_{\pi P}} \) are all positive. In all but one case the hypothesis that the ratio of parameters is equal to one can be rejected at the 5% level in favor of the coefficient being less than 1.0. Quarterly US data is the only case that fails to reject the hypothesis. A finding of less than one-for-one nominal rate responses to inflation in 18 of the 19 cases is a highly robust result. But is this overwhelming refutation of Fisher simply a consequence of endogenous money growth?

The long-run nominal interest rate response to a 1 percentage point permanent increase in the inflation rate is related to structural parameters based on Equation (36):

\[
\frac{R_{iP}}{R_{\pi P}} = \frac{\eta \alpha_{\pi \mu}^2 + \eta \alpha_{\pi \rho}^2 + \alpha_{\pi \rho} \alpha_{r \rho} + \eta \alpha_{\pi \lambda}^2}{\alpha_{\pi \mu}^2 + \alpha_{\pi \rho}^2 + \alpha_{\pi \lambda}^2} = \eta + \frac{\alpha_{\pi \rho} \alpha_{r \rho}}{\alpha_{\pi \mu}^2 + \alpha_{\pi \rho}^2 + \alpha_{\pi \lambda}^2} > \eta ,
\]

where the inequality is obtained from Assumptions S6 and S7. Recall that \( \eta \) is the response of the nominal rate to a permanent increase in inflation. This means that if money growth is endogenous to savings shocks, the Rapach (2003) estimate will be an upward biased estimate of the long-run response of nominal rates to inflation. In other words, the structural effect of inflation on nominal rates is on-average smaller than what the model estimates. Therefore, Rapach (2003)’s rejection of the Fisher Effect in favor of a less than one-for-one reaction of nominal rates to inflation is robust to endogenous money growth.

---

\(^{42}\)Appendix A provides a more formal justification for each assumption.

\(^{43}\)Neyapti (2003) and Sill (2005) provide evidence that S6 may be a good assumption for less-developed economies with central banks that are less independent. They also suggest that \( \alpha_{\pi \rho} = 0 \) may be warranted for economies that are developed or have independent central banks. Combining this restriction while maintaining all other structural assumptions does not change the conclusions of this paper. None of the results are qualitatively altered, except that now the Fisher Effect test becomes unbiased.
6.2 Positive Output Effects are Robust to Endogenous Money Growth

Rapach (2003)’s quarterly estimates of $R_{yP}/R_{\pi P}$ all are positive but not significantly different from zero at the 5% level, and with annual data 12 of the 14 estimates are not significantly different from zero. In 9 of those countries positive estimates are obtained while the other 3 are not. Only Austria and The Netherlands obtain significant estimates, and in both cases they are positive. The positive estimates for Belgium, France, and Ireland are each very nearly significant. The US estimate is negative as is Denmark’s, while Italy’s estimate is essentially zero.

Bullard and Keating (1995) and Rapach (2003) estimate models using somewhat different samples of countries, although 6 countries are common to both: Australia, Austria, Germany, Ireland, the United Kingdom, and the United States. Bullard and Keating (1995) used annual data, and so comparing those results with Rapach (2003)’s annual results may be useful. Also the papers differ in terms of sample periods, unit root testing procedures, and methods for selecting VAR lag lengths. Despite these variations, the results for superneutrality are qualitatively quite similar. For the 6 common countries the estimates are all positive, except for the United States. The main difference is that estimates for Germany and the UK are statistically significant in Bullard and Keating (1995) but not so in Rapach (2003).

A likely reason for some differences in results is variation in model size. Even if lag length remains the same, a trivariate VAR system estimates a lot more coefficients than a bivariate model. For example, both papers chose a lag length of one for Germany, but that means 9 lag coefficients in Rapach (2003)’s trivariate VAR compared to only 4 in Bullard and Keating (1995)’s bivariate framework. With more than twice as many coefficients, the trivariate model will have greater sampling error for a fixed sample size, and note that both papers use approximately the same sample for Germany. The difference in the number of coefficients is even larger for the UK. The lag length in Rapach (2003) was 4 while in Bullard and Keating (1995) it was 2. Thus the trivariate model estimated 36 lag coefficients whereas the bivariate model estimated only 8.

While there is some variation across the two papers, overall the results from Rapach (2003)’s trivariate model are fairly similar to those in the bivariate model of Bullard and Keating (1995). Each obtains some significant positive estimates, a handful of nearly significant positive estimates, and the majority of countries obtain positive estimates even though a large fraction of those is insignificant. This raises the question: Does downward bias also exist in the trivariate model?

Equation (36) shows how the long-run output effect of a permanent 1 percentage point shock to inflation is mapped into the structure:

$$
\frac{R_{yP}}{R_{\pi P}} = \frac{\alpha_{\pi\mu} \alpha_{y\mu} + \alpha_{\pi\rho} \alpha_{y\rho} + \alpha_{\pi\lambda} \alpha_{y\lambda}}{\alpha_{\pi\mu}^2 + \alpha_{\pi\rho}^2 + \alpha_{\pi\lambda}^2} = \frac{\alpha_{y\mu}}{\alpha_{\pi\mu}} + \frac{\alpha_{\pi\rho} \alpha_{y\rho}}{\alpha_{\pi\mu}^2} + \frac{\alpha_{\pi\lambda} \alpha_{y\lambda}}{\alpha_{\pi\mu}^2},
$$

Equation (36) shows how the long-run output effect of a permanent 1 percentage point shock to inflation is mapped into the structure:

44Rapach (2003)’s US estimate with quarterly data is positive, although it is not significant.

45Rapach (2003) had fewer estimates that were zero or negative, about 21% (3 out of 14) of the countries in that sample compared with about 44% (7 of 16) in Bullard and Keating (1995). This may be explained by the fact that all the countries in Rapach (2003) had average inflation less than 10% while Bullard and Keating (1995)’s sample included 6 countries with inflation rates of 15% or more and in some cases much more.
with the last expression obtained by dividing by $\alpha_{\pi\mu}^2$, which is the same operation used earlier. The first term in the numerator is again what we’d like to estimate. Now there are two other terms in the numerator and two in the denominator that interfere with estimating this structural effect of interest. In this case the denominator is greater than 1 if money growth is endogenous to either technology or savings shocks. And based on the structural assumptions, the second and third terms in the numerator are both negative. The same result as before is obtained for this larger model: if $\frac{\alpha_{\pi\mu}}{\alpha_{\pi\mu}} > 0$, the estimate is biased downward.

In general, the estimate is downward biased when:

$$\frac{R_y}{R_{\pi}} < \frac{\alpha_{\pi\mu}}{\alpha_{\pi\mu}}. \quad (39)$$

Plugging Equation (38) into (39) and simplifying yields:

$$\frac{\alpha_{\pi\mu}}{\alpha_{\pi\mu}} > \frac{\alpha_{\pi\rho} \alpha_{\pi\rho} + \alpha_{\pi\lambda} \alpha_{\pi\lambda}}{\alpha_{\pi\rho}^2 + \alpha_{\pi\lambda}^2}. \quad (40)$$

The right hand side of this expression is negative given the structural assumptions. This means the estimate is downward biased if the long-run output response to an exogenous increase in money growth is positive, zero, or even for some negative values of this effect.

Given the results for real interest rates, Rapach (2003)’s two cases in which the estimate is positive and statistically significant are evidence of Mundell-Tobin Effects. Bias from endogenous money growth means that the actual Mundell-Tobin effects are likely to be even larger. Most of the long-run output effects are positive. But most are statistically insignificant, possibly a consequence of downward bias. If superneutrality or a reverse Mundell-Tobin effect (whereby in the long-run an exogenous increase in money growth yields higher real rates and lower output) were the norm and money growth was endogenous, one would have expected estimates to be mostly negative. The fact that they are mostly positive is consistent with Bullard and Keating (1995)’s finding that estimates from low inflation countries were overwhelmingly positive and Keating et al. (2021)’s interpretation of those results as providing evidence of Mundell-Tobin Effects. And as discussed earlier, greater sampling variation from the larger VARs of Rapach (2003) likely plays a role in the trivariate models finding fewer countries with significant estimates.

7. Conclusion

There are various reasons money growth could respond to shocks to the real economy. Concern that such a response may invalidate conclusions about money growth and inflation from SVAR studies is warranted. This paper provides a careful accounting of endogenous money growth. Interpretations of empirical results from structural VARs as evidence that higher money growth raises real output and lowers real interest rates in the long run are shown robust to endogeneity, particularly for low inflation countries. Findings for some high inflation countries provide support for the hypothesis that money growth was endogenous.

This paper arrives at qualitative conclusions. These are important since theory alone is unclear about even the sign of inflation’s long-run effect on either output or the real interest rate.
Narrowing the range of these non-neutral effects is an important direction for future research. Additional work to estimate the effects of money growth and inflation on real variables along the lines of Keating et al. (2021) may prove useful.

A more precise quantitative understanding of these effects is important. Conventional wisdom about the optimal inflation target is typically derived from economic models that do not permit a Mundell-Tobin effect. One argument for aiming for an inflation rate of around 2 percent is to serve as a buffer against deflation. Another one is that nominal rates rise with inflation, giving monetary policy more room to reduce rates before reaching an effective lower bound. But if somewhat higher inflation in the long run raises output and lowers the real interest rate, a higher target for inflation than most central banks have lately been willing to consider might be a better strategy.

Appendix A - Description and Summary Statistics of Inflation Measures

This appendix derives a structure in which Rapach (2003)’s preference shock is reinterpreted as an exogenous shock to savings. The problem for a preference shock interpretation is that when this shock has a unit root it drops out of the standard Euler equation and thus has no permanent real interest rate effect.

Variables are in logarithmic form except for the nominal interest rate, the real rate, and the inflation rate. All variables are first-differenced to accord with Rapach (2003)’s empirical model. Note that differencing removes constants that would naturally appear in some of these structural equations. For example, prior to taking the first difference the log of the demand for capital has a constant that depends on $\alpha$. All parameters take on positive values.

Assume a standard Cobb-Douglas function:

$$\Delta y = \alpha \Delta k + (1 - \alpha)(\Delta n + \Delta \tau)$$

which determines real output on the basis of capital ($k$), labor ($n$), labor augmenting technology ($\tau$), and a parameter that is subject to the usual boundaries: $0 < \alpha < 1.0$.

Profit is maximized subject to this production function. Assuming perfectly competitive factor markets yields a standard capital demand function from the first order condition which equates the marginal product of capital to its marginal cost (both in logarithms and then first differenced):

$$\Delta(y - k) = \phi \Delta r.$$  \hspace{1cm} (A.2)

The cost of capital is assumed to be an increasing function of the real interest rate. Combining the production function with the demand for capital yields:

$$\Delta y = \left(\frac{-\alpha \phi}{1 - \alpha}\right) \Delta r + \lambda^*.$$

$\lambda^* = \Delta n + \Delta \tau$ are supply shocks generated by both technology and labor supply, assuming the long-run labor supply curve is inelastic. Note that a qualitatively similar result would obtain if labor supply is assumed a function of the real wage, the real interest rate, and an exogenous
labor supply shock. This is seen by combining that equation with the standard labor demand curve derived from the Cobb-Douglas production function and then substituting the solution for equilibrium $\Delta n$ into the production function. In that case the coefficient on $\Delta r$ would be negative under plausible assumptions about the structural parameters, and $\lambda^*$ will still be a linear function of exogenous shocks to labor supply and technology.

Next assume desired savings is an increasing function of output, decreasing in real interest rate, and $\gamma^*$ is an exogenous shock associated with government savings:

$$\Delta s = \Delta y - \omega \Delta r + \gamma^*.$$  \hfill (A.4)

Note that long-run income elasticities of capital demand and desired savings are assumed to be the same. If these elasticities are not identical, then as $y$ rises, $r$ would trend either upward or downward depending on which income elasticity is larger. This is ruled out based on Yi and Zhang (2017), which found by examining roughly the last 60 years of real interest rate data for 20 modern economies that there is no trend in real rates. The real rate does exhibit long swings, with periods when it is persistently high or persistently low. This evidence suggests the real rate has a unit root, and the statistical model allows for that.\footnote{The evidence suggests if the real rate has a unit root it is without drift, while if output has a unit root it will require drift.}

In the long run, $\Delta s = \Delta I = \Delta k$, where $I$ is the logarithm of investment. The first equality comes from national income accounts which equate savings to investment. The second equality can be derived from the standard capital accumulation equation, which equates gross investment to net investment plus depreciation. Assuming the depreciation rate and the growth rate of aggregate capital are constant in the long-run, these constants disappear after taking logarithms and differencing. From the previous equation, set desired savings equal to the demand for capital and solve for the real rate:

$$\Delta r = \frac{-\gamma^*}{\omega + \phi} = \rho^*.$$  \hfill (A.5)

which justifies redefining Rapach (2003)'s preference shock as a negative savings shock.

Rapach (2003) assumed permanent changes in inflation were exogenous and associated with permanent shocks to money growth:

$$\Delta \pi = \mu^*.$$  \hfill (A.6)

The final structural equation under the null hypothesis is the Fisher Equation:

$$\Delta i = \Delta r + \Delta \pi.$$  \hfill (A.7)

Now write each structural shock as a product of its standard deviation and an iid unit variance shock:

$$\epsilon_t^* = \sigma_t \epsilon_t \quad \text{for} \quad \epsilon = \mu, \rho, \lambda \quad \text{with} \quad \mathbb{E} \epsilon^2 = 1,$$  \hfill (A.8)

where $\sigma_t > 0$ is the standard deviation for shock $\epsilon$.

Combining these equations for nominal and real interest rates, inflation, and output, yields a structure that captures Rapach (2003)'s null hypothesis. If variables are ordered: $X_t = (\pi, i, y)'$ and shocks ordered: $\epsilon_t = (\mu, \rho, \lambda)'$, then the matrix of long-run effects will be:
where the alpha parameters are functions of structural parameters and the standard error of the shock relevant to a particular column. This is a restricted lower triangular matrix. It remains lower triangular if there is non-superneutrality, \( \alpha_{y\mu} \neq 0 \), or failure of the Fisher Effect, \( \alpha_{i\mu} \neq \alpha_{\pi\mu} \).

How does endogenous money growth affect those interpretations? Allowing long-run money growth to potentially respond endogenously to savings and technology shocks:

\[
\Delta \pi = \alpha_{\pi\mu} \mu + \alpha_{\pi\rho} \rho + \alpha_{\pi\lambda} \lambda.
\] (A.10)

If superneutrality holds as well as the Fisher effect, then nominal interest rates move one-for-one with inflation and the long-run structure becomes:

\[
\theta(1) = \begin{bmatrix}
\alpha_{\pi\mu} & 0 & 0 \\
\alpha_{\pi\mu} & \alpha_{\pi\rho} & 0 \\
0 & \alpha_{y\rho} & \alpha_{y\lambda}
\end{bmatrix},
\] (A.9)

\[
\theta(1) = \begin{bmatrix}
\alpha_{\pi\mu} & \alpha_{\pi\rho} & \alpha_{\pi\lambda} \\
\alpha_{\pi\mu} & \alpha_{\pi\rho} + \alpha_{\pi\rho} & \alpha_{\pi\lambda} \\
0 & \alpha_{y\rho} & \alpha_{y\lambda}
\end{bmatrix}.
\] (A.11)

The structure is no longer recursive under the null and Rapach (2003)’s tests would no longer work as intended.

The final step is to allow for possible failure of the two neutrality propositions. First, the economy is not restricted to be superneutral with respect to output: \( \alpha_{y\mu} \neq 0 \). Second, allow the nominal interest rate to possibly not respond one-for-one to a permanent movement in inflation, which means replacing (A.7) by:

\[
\Delta i = \Delta r + \eta \Delta \pi 
\] (A.12)

and \( \eta \) is restricted (although a negative value is implausible). Equation (35) in the text is obtained by combining these two assumptions with (A.11).
References


Canova, F., and DeNicoló, G. (2002). Monetary disturbances matter for business fluctuations in...
Implications of Endogenous Money Growth for Some Tests of Superneutrality and...


