An Alternative Explanation of the Price Puzzle

Paolo Giordani*

January 2001

Abstract

This paper proposes a simple explanation for the frequent appearance of a price puzzle in VARs designed for monetary policy analysis. It suggests that the best method of solving the puzzle implies a close connection between theory and empirics rather than the introduction of a commodity price. It proves that the omission of a measure of output gap (or potential output) spuriously produces a price puzzle in a wide class of commonly used models. This can happen even if the model admits a triangular identification and if the forecasts produced by the misspecified VAR are optimal. In the framework of a model due to Svensson, the omission of a measure of output gap is shown to generate several other incorrect conclusions. When the model is tested on US data, all predictions are supported.

Keywords: VAR, monetary policy, misspecification, output gap, potential output, technology shocks.

JEL: E 30, E 52.

*Stockholm School of Economics. e-mail address nepgi@hhs.se. I would like to thank David Domeij, Martin Eichenbaum, Lars Ljungqvist, and Anders Vredin for comments, and Charles Evans for providing data. A special thank to Paul Söderlind and Lars Svensson for comments and discussions. Financial help from the Wallanders Foundation is gratefully acknowledged.

1 Introduction

A vast literature has produced a reference framework for VAR analysis of monetary policy.[1] This reference VAR includes a commodity price index. The first VAR studies showed that omitting a commodity price and taking a short interest rate as the policy instrument produced a response of the price level to contractionary monetary policy shocks which was positive for many quarters, a finding that took the name of price puzzle. Sims (1992) proposed a rationale for the puzzle, and a way to fix it. His conjecture was that the information set available to policy makers may include variables useful in forecasting future inflation that the econometrician has not considered. If the VAR forecast of inflation is in fact a poor one, the VAR will mistakenly identify as shocks movements in the instrument of policy which are in fact endogenous responses to signals of future inflation, hence the finding that prices increase after a contractionary monetary policy shock.[2] Sims himself (1992) and later studies building on this suggestion have found that the puzzle disappears in the US, at least to a large extent, when the VAR is extended to include a commodity price index, a variable useful in forecasting inflation.

Besides solving the price puzzle, the inclusion of a commodity price changes the overall picture of monetary policy, in that the response of output to a MP shock is smaller and MP shocks are less important in the variance decomposition of output and of the federal funds rate (the policy instrument). Based on these results, Leeper, Sims and Zha (1996) warn that the exclusion of a commodity price can result in serious misspecification.

But while no one wants a price puzzle in their VAR, eliminating it sometimes comes at a cost since the models monetary economists work with do not include

[1] For a summary of this literature see Christiano, Eichenbaum and Evans (1998) or Leeper, Sims and Zha (1996). For a thorough presentation of the framework VAR for monetary policy analysis, see Favero (2000).

[2] Henceforth MP shock.

a commodity price. Of course the models are not meant to be complete representations of reality, and if monetary authorities do react to information not incorporated in the models, so much worse for the models. Nevertheless, having to include a commodity price in the VAR can be disturbing if a researcher is trying to bring a model to the data and she is interested in identifying all shocks, or anyway more than just MP shocks. For example, how are we to interpret the structural shocks if the VAR has two price levels, say CPI and commodity prices, but the theoretical model only has one? Adding variables to the VAR to solve the price puzzle makes interpretation and identification of shocks other than the MP shock more problematic and less model-driven. This paper argues that this situation may be avoidable.

The paper explores the possibility that the price puzzle may be due to something other than the omission of a variable useful in forecasting inflation (such as a commodity price). It shows that a wide class of models produces a price puzzle when subjected to a seemingly innocent misspecification common in applied research: output is used in applications while theory speaks of the output gap. The key requirements needed to produce a puzzle is that the monetary authority has the output gap in its policy function and that there are lags in the transmission of monetary policy, so that monetary policy affects output first and then inflation. The intuition is that since the output gap is omitted from the inflation equation, the interest rate spuriously appears in the equation with a positive coefficient, because the interest rate reacts positively to output gap increases.

The rest of the paper proceeds as follows. Section 2 presents a model for monetary policy analysis (Svensson (1997)), which incorporates in a simple form some key features of more complex models and admits a triangular identification scheme in the order: potential output (or output gap), output, inflation and interest rate (the model also admits a three variable representation including output gap and omitting output). Taking the Svensson model as the data generating process (DGP), this section explores analytically the consequences of estimating a three variable VAR that includes output but not the output gap. Among other things, a price puzzle emerges and the variance of the MP shocks is overestimated. The impact of a MP shock on output is also overestimated. The consequences of the misspecification are also shown through impulse responses, giving more color to the analytical results. These results extend to a rich class of models (Section 2.3).

Section 3 takes the theory to the data, using as output gap a measure of capacity utilization produced by the Federal Reserve Board. A three variable VAR in the order: output gap, inflation and federal funds rate, is compared to a VAR including output rather than the output gap. The second VAR produces a large price puzzle, the first none. In the first VAR monetary policy is more endogenous and accounts for much less of the forecast error variance of output. Overall, the results produced by the first VAR are closer to those implied by theory and by larger VARs that include a commodity price. Section 3.1 estimates a four variable VAR (derived from the model) in the order: potential output, output, inflation, federal funds rate, allowing technology shocks to enter the picture. Technology shocks identified with short run restrictions taken from the model yield predictions consistent with the model and with the assumption (not imposed) that only technology shocks affect output in the long run. Section 4 argues that the commodity price index does not solve the price puzzle because it is useful in forecasting inflation, but rather because it is correlated with the output gap. Section 5 concludes.

2 A simple model for monetary policy analysis: Svensson (1997)

Svensson (1997) presents a model designed to capture some key features of the transmission mechanism of monetary policy. In fact, it is more generally a model of business cycle fluctuactions. The same model is used in Rudebusch and Svensson (1999), in Judd and Rudebusch (1998) and, extended to a small open economy, in Ball (1999). A forward looking version appears in Clarida, Gali and Gertler (1999) and in Svensson (2000a and 2000b). Romer (2000) presents the same model as an improvement over the traditional IS-LM. The model consists of an IS equation, a Phillips curve and a Taylor rule obtained from the monetary authority's optimization problem. This core three-equation structure is shared by many recent New Keynesians models for monetary policy analysis. A distinct feature of the model is that it incorporates delays in the transmission of monetary policy. Monetary policy can affect output only with a lag. Output, in turn, affects inflation with a lag. Since the transmission from policy action to prices goes through output variations, monetary policy affects prices with two lags. The IS and Phillips relations are backward-looking, but section 2.3 shows that the main results hold in a very general framework, which includes forward looking and (totally or partially) microfounded models, the key requirement being that they display delays in the transmission of monetary policy. The IS relation is given by

$$y^g_{t+1} = \beta_y y^g_t - \beta_r (i_t - \pi_t) + \epsilon^{AD}_{t+1}, \tag{1}$$

where i_t is a short term interest rate set by the monetary authority, y^g is the output gap, defined as $y^g_t = Y_t - Y^N_t$, where Y_t is the log of output and Y^N_t the log of natural (or "potential") output. Natural output is assumed to follow an exogenous AR(1) process[3]

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

The Phillips curve is modelled as

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

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

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

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

$$i_t = \gamma_\pi \pi_t + \gamma_y y_t^g. \tag{5}$$

A monetary policy shock can be added (ad hoc) by supposing that the Taylor rule is not followed deterministically. In that case, the shock ϵ^{MP} with std σ_{MP} is added to the Taylor rule.

2.1 The correct identification

AD shocks affect output but not inflation contemporaneously while MP shocks affect neither output nor inflation contemporaneously. Technology shocks increase output contemporaneously, leaving output gap, inflation and interest rate

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

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

unchanged (the extension of Section 2.3 can accomodate technology shocks that affect all variables). The model delivers a three variable VAR with triangular identification in the order: output gap, inflation, interest rate (output gap and inflation can be reversed). Four variable formulations are also admissible, with any two of output, output gap, natural output, appropriately ordered. However, the model does not justify a three variable VAR including output, inflation and interest rate, which is the core of VARs that researchers have estimated in practice.[5]

2.2 From the VAR implied by theory to the empirical VAR, taking a false step

Let the *DGP* be given by equations (1) − (5). The Taylor rule is assumed deterministic for simplicity (all results generalize to the case $\sigma_{MP} > 0$: the Appendix shows the form taken by the system if $\sigma_{MP} > 0$, and a simulation for this case appears later in this section). Suppose that a researcher estimates a VAR including: output, inflation and interest rate (identified in the same order) but not the output gap.[6] What are the effects of this common and seemingly innocent change?

I start investigating the consequences of this misspecification by relating the true structural moving average representation to the one recovered by the misspecified VAR. To avoid mixing problems of misspecification and of parameter uncertainty due to small sample size, assume that the sample is large. The *DGP* has a VAR(1) representation

$$A_0 X_t = A_1 X_{t-1} + \epsilon_t,$$

[5] Two exceptions are Leichter and Walsh (1999) and Rudebusch and Svensson (1999).

[6] This group of three variables, with this same ordering, plus a commodity price index ordered after prices, is the core of the framework VAR model for monetary policy analysis (see, for example, Bagliano and Favero (1998) and Favero (2000)).

which can be inverted to obtain the moving average representation of the DGP

$$X_t = \sum_{i=0}^{\infty} D_i \epsilon_{t-i},$$

where

$$\begin{aligned} X_t^{'} &= \{Y_t^N, Y_t, \pi_t, i_t\} \qquad \epsilon_t^{'} = \{\frac{\epsilon_t^N}{\sigma_N}, \frac{\epsilon_t^{AD}}{\sigma_{AD}}, \frac{\epsilon_t^{CP}}{\sigma_{CP}}\} \qquad VCV(\epsilon_t) = I \\ D_0 &= \begin{bmatrix} \sigma_N & 0 & 0 \\ \sigma_N & \sigma_{AD} & 0 \\ 0 & 0 & \sigma_{CP} \\ 0 & \gamma_y \sigma_{AD} & \gamma_\pi \sigma_{CP} \end{bmatrix} \end{aligned} \tag{6}$$

The researcher is estimating a VAR in Y, π, i[7], working under the (erroneous) assumption that the moving average representation for the structural residuals is given by (asterisks denote the misspecified system)

$$Z_t = \sum_{i=0}^{\infty} D_i^* \epsilon_{t-i+1}^*,$$

where $Z_t^{'} = \{Y_t,\ \pi_t,\ i_t\}$. The researcher identifies the system by assuming that D_0^* has a lower triangular structure (the system of equations is assumed recursive). Since $D_0^* D_0^{*\prime} = \sum$, where $\sum$ is the variance-covariance matrix of the reduced form residuals, the recursive assumption implies that $D_0^* = Cholesky(\sum)$. The researcher will then interpret D_0^* as

$$D_0^* = \begin{bmatrix} \sigma_{AD}^* & 0 & 0 \\ b_{21} & \sigma_{CP}^* & 0 \\ \gamma_y^* \sigma_{AD}^* & \gamma_\pi^* \sigma_{CP}^* & \sigma_{MP}^* \end{bmatrix}$$

If $\sigma_{MP} = 0$, the exclusion of Y^N produces no loss of fit in any equation, a result driven by the deterministic form of the Taylor rule that allows to retrieve Y^N (and thus y^g) with no error as a linear combination of the three variables of the system. Therefore the misspecified VAR produces optimal forecasts of inflation

[7] Throughout the paper I refer to the VAR in output, inflation and interest rate as "misspecified VAR".

at all time horizons, excluding the possibility that the standard explanation for the price puzzle may be relevant in this setting. Consequently

$$\sum = D_0^* D_0^{*'} = [D_0 D_0']_2^4,$$

where $[D_0 D_0^{'}]_2^4$ is the 3×3 matrix obtained deleting the first row and the first column of $D_0 D_0^{'}$.

$$\begin{aligned}
[D_0 D_0^{'}]_2^4 &= \begin{bmatrix} \sigma_N^2 + \sigma_{AD}^2 & 0 & \gamma_y \sigma_{AD}^2 \\ & \sigma_{CP}^2 & \gamma_\pi \sigma_{CP}^2 \\ & & \gamma_y^2 \sigma_{AD}^2 + \gamma_\pi^2 \sigma_{CP}^2 \end{bmatrix} = D_0^* D_0^{*'} = \qquad (7) \\
&= \begin{bmatrix} \sigma_{AD}^{*2} & \sigma_{AD}^* b_{21} & \gamma_y^* \sigma_{AD}^{*2} \\ & \sigma_{CP}^{*2} + b_{21}^2 & b_{21} \gamma_y^* \sigma_{AD}^* + \gamma_\pi^* \sigma_{CP}^{*2} \\ & & \gamma_y^{*2} \sigma_{AD}^{*2} + \gamma_\pi^{*2} \sigma_{CP}^{*2} + \sigma_{MP}^{*2} \end{bmatrix}
\end{aligned}$$

The relations between the actual and estimated shocks are straightforwardly obtained from the equalities in (7). b_{21} is correctly set to zero, since $\sigma_{AD}^* b_{21} = 0$. It follows that:

1. $\sigma_{AD}^{*2} = \sigma_N^2 + \sigma_{AD}^2$. **The variance of the labelled AD shock is the sum of the variances of the AD shock and of the technology shock** The importance of AD shocks in the variance decomposition of output is overestimated.

2. $\sigma_{AD}^{*2} = \sigma_{AD}^2 + \sigma_N^2$ and $\gamma_y \sigma_{AD}^2 = \gamma_y^* \sigma_{AD}^{*2}$ imply $\frac{\gamma_y}{\gamma_y^*} = \frac{\sigma_{AD}^2 + \sigma_N^2}{\sigma_{AD}^2} > 1$, that is the output gap coefficient in the Taylor rule is underestimated. Some simple algebra gives $\gamma_y \sigma_{AD} - \gamma_y^* \sigma_{AD}^* = \gamma_y \sigma_{AD} (1 - \frac{\sigma_{AD}}{\sqrt{\sigma_{AD}^2 + \sigma_N^2}}) > 0$. This means that **the intensity of the response of the monetary authority to a one std AD shock is underestimated** even though the std of AD shocks is overestimated. The intuition is that some unforecasted movements in output are due to technology shocks, to which monetary policy does not respond. Since the misspecified VAR registers a small average reaction of the interest rate to unforecasted output movements, the

coefficient γ_y in the Taylor rule is underestimated. The underestimation grows with σ_N.

3. $\sigma_{CP}^2 = \sigma_{CP}^{*2}$, following from the fact that $b_{21} = 0$.

4. Finally, **the variance of MP shocks is overestimated**. To derive the result, start from (7), which sets $\gamma_y^2 \sigma_{AD}^2 + \gamma_\pi^2 \sigma_{CP}^2 = \gamma_y^{*2} \sigma_{AD}^{*2} + \gamma_\pi^{*2} \sigma_{CP}^{*2} + \sigma_{MP}^{*2}$. Use the results obtained so far, namely

(a) $\gamma_\pi^2 \sigma_{CP}^2 = \gamma_\pi^{*2} \sigma_{CP}^{*2}$

(b) $\sigma_{AD}^{*2} = \sigma_N^2 + \sigma_{AD}^2$

(c) $\frac{\gamma_y}{\gamma_y^*} = \frac{\sigma_{AD}^2 + \sigma_N^2}{\sigma_{AD}^2}$

to obtain

$$\sigma_{MP}^{*2} = \gamma_y^2 \frac{\sigma_{AD}^2 \sigma_N^2}{\sigma_{AD}^2 + \sigma_N^2} > 0. \tag{8}$$

Even though the Taylor rule is deterministic, the VAR finds that the variance of the labelled MP shock is strictly positive. The intuition is that since the interest rate does not react in the same way to technology and AD shocks, when a movement in output (of a given amount) is observed the VAR will sometimes register a certain change in the interest rate (when the movement is caused by an AD shock) and sometimes a different change (when caused by a technology shock) and will be tricked into interpreting this as random behavior of the monetary authority.[8] If $\sigma_{MP} > 0$, the right-end-side of equation (8) gives a lower bound for $\sigma_{MP}^{*2} - \sigma_{MP}^2$. Notice that if $\sigma_N = 0$, all misspecifications disappear, as they should since in that case $y_t^g = Y_t$.

I now use the result that $\sigma_{MP}^* > 0$ to prove that the misspecified system will display a price puzzle. The strategy is to derive the coefficients of the inflation equation in the misspecified VAR.

[8] In this model the interest rate does not react at all to technology shocks, but the intuition applies more generally, as long as the monetary authority reacts differently to technology and AD shocks.

Use the Taylor rule in (5) to obtain an expression for the output gap

$$y_t^g = \frac{1}{\gamma_y}(i_t - \gamma_\pi \pi_t), \tag{9}$$

and substitute it into the Phillips relation. This yields

$$\pi_{t+1} = [1 - \alpha_y \frac{\gamma_\pi}{\gamma_y}]\pi_t + \frac{\alpha_y}{\gamma_y} i_t + \epsilon_{t+1}^{CP}. \tag{10}$$

Equation (10) is both the inflation equation in the VAR and the structural equation of the recursive system, since $b_{21} = 0$ (see (7)) implies that Y_{t+1} has a zero coefficient. If $\sigma_N = 0$, it follows that $y_t^g = Y_t$, so there is no misspecification and (10) is equivalent to

$$\pi_{t+1} = \pi_t + \alpha_y Y_t + \epsilon_{t+1}^{CP}. \tag{11}$$

Y_t, π_t, and i_t are perfectly collinear and choosing between (10) and (11) is a matter of taste. On the other hand, if $\sigma_N > 0$, no other autoregressive representation fits as well as (10). Therefore OLS will retrieve (10). The reason why i_t appears with a positive coefficient in (10) is that movements in the interest rate help retrieve movements in the output gap, which is omitted. Since $\alpha_y/\gamma_y > 0$ and output does not appear in the equation, the impact of a MP shock (which causes i_t to be higher than forecasted) on inflation is estimated to be zero contemporaneously and positive at one lag. In other words, a positive response of inflation to a contractionary MP shock (a price puzzle) at lag one is guaranteed as long as the variance of the retrieved MP shocks is estimated to be strictly positive, which will be the case if $\sigma_N > 0$ (see equation (8)). The magnitude of the puzzle at lag one is given by $\frac{\alpha_y}{\gamma_y}\sigma_{MP}^*$, so it grows with the variance of technology shocks (see equation (8)).

To gain further understanding of the puzzle, it is useful to show that the misspecified MP^* shocks are positively correlated with the true AD shocks and negatively correlated with the true technology shocks.

In the DGP, the one-step-ahead forecast error is given by

$$i_{t+1} - E_t i_{t+1} = \gamma_y \epsilon_{t+1}^{AD} + \gamma_\pi \epsilon_{t+1}^{CP}, \tag{12}$$

while the one-step-ahead forecast error in the misspecified model is given by

$$i_{t+1} - E_t^* i_{t+1} = \gamma_y^* \epsilon_{t+1}^{*AD} + \gamma_\pi^* \epsilon_{t+1}^{*CP} + \epsilon_{t+1}^{*MP}. \tag{13}$$

Asterisks denote the shocks obtained from the misspecified VAR. Since the assumption that $\sigma_{MP} = 0$ implies $E_t i_{t+1} = E_t^* i_{t+1}$, we can equate the right-hand-sides and use the results just obtained, namely $\gamma_\pi \epsilon_{t+1}^{CP} = \gamma_\pi^* \epsilon_{t+1}^{*CP}$, $\epsilon_{t+1}^{*AD} = \epsilon_{t+1}^{AD} + \epsilon_{t+1}^{N}$ to obtain as expression for ϵ^{*MP},

$$\epsilon_{t+1}^{*MP} = (\gamma_y - \gamma_y^*)\epsilon_{t+1}^{AD} - \gamma_y^* \epsilon_{t+1}^{N}. \tag{14}$$

Since $cov(\epsilon_{t+1}^{AD}, \epsilon_{t+1}^{N}) = 0$ by assumption, using $\frac{\gamma_y}{\gamma_y^*} = \frac{\sigma_{AD}^2 + \sigma_N^2}{\sigma_{AD}^2} > 0$ gives

$$cov(\epsilon_{t+1}^{*MP}, \epsilon_{t+1}^{AD}) = (\gamma_y - \gamma_y^*)\sigma_{AD}^2 > 0, \tag{15}$$

$$cov(\epsilon_{t+1}^{*MP}, \epsilon_{t+1}^{N}) = -\gamma_y^* \sigma_N^2 < 0. \tag{16}$$

These results provide further intuition for the origin of the price puzzle: the misspecified monetary policy shocks are positively correlated with the true aggregate demand shock, which in turn raise inflation with a lag. Since at lag one the true monetary policy shocks cannot affect inflation, only the spurious part is active at lag one, so we are certain to find a price puzzle. Moreover, monetary policy shocks are spuriously correlated with technology shocks. This means that the misspecified impulse response of output to a contractionary monetary policy shock is contaminated by the response of output to a negative technology shock. If potential output is more persistent than the output gap, the response of output to a monetary policy shock will be longer lived than the true one.

There are more potentially erroneous conclusions that can be derived from the misspecified system. To illustrate them, I obtain the reduced form equation for output in the misspecified system. Start with the IS equation (1) and eliminate y^g_{t+1}, y^g_t using the definition, $y^g_t = Y_t - Y^N_t$. Move Y^N_{t+1} to the right-hand-side and substitute it using (2). This leaves ϵ^N_{t+1} and Y^N_t on the right-hand-side. Finally, eliminate Y^N_t from the right-hand-side by rearranging (9) as

$$Y^N_t = Y_t - \frac{1}{\gamma_y}(i_t + \gamma_\pi \pi_t), \tag{17}$$

which gives the equation for output in the misspecified model

$$Y_{t+1} = \rho Y_t + [(\rho - \beta_y)\frac{\gamma_\pi}{\gamma_y} + \beta_r]\pi_t - [(\rho - \beta_y)\frac{1}{\gamma_y} + \beta_r]i_t + \epsilon^{AD}_{t+1} + \epsilon^N_{t+1}. \tag{18}$$

Some implications of equation (18) are worth noticing.

- If $\rho > \beta_y$, for example when the output gap is stationary while technology has a unit root, AD shocks will appear to have a more persistent effect on output than they actually do.

- If $\rho > \beta_y$, the effect of a given interest rate shock on output one step ahead is overestimated, since the coefficient attached to i_t is larger than β_r. The same is true of CP shocks.

The derivation of (10) and (18) assumed that the Taylor rule is deterministic. The Appendix shows the form taken by the system if $\sigma_{MP} > 0$, and a simulation for this case appears later in this section. If $\sigma_{MP} > 0$, the misspecified system (Y, π, i) is no longer VAR(1), but VARMA(2,1). This implies that the econometrician who is selecting lag length for a VAR is likely to choose a VAR with more than one lag, and will produce sub-optimal fit and forecasts even is she estimates a VARMA(2,1).

A more complete picture of the consequences of the misspecification can be gained from looking at impulse responses. The experiment is as follows. Each graph plots the response of output or inflation or interest rate to a shock in the theoretical economy together with the response to the same shock in the misspecified three variable VAR (output, inflation, interest rate).

The model parameters are set as in Ball (1999): $\alpha_y = 0.4$, $\beta_y = 0.8$, $\beta_r = 1$. Reflecting the idea that potential output is a highly persistent process, I set $\rho = 0.98$. The standard deviations are $\sigma_{AD} = \sigma_{CP} = \sigma_N = 1$. For ease of comparison, the parameters in the Taylor function are not set to the optimal value in each case, but are kept constant at $\gamma_y = 0.5$, $\gamma_\pi = 1.5$.[9] These parameters are kept fixed. The only difference between Figure 1 and Figure 2 is the standard deviation of the true monetary policy shock. In Figure 1 $\sigma_{MP} = 0$, so all results are analytical. Recall that the differences between theoretical and misspecified responses cannot be accounted for by parameter uncertainty. The response of output to a labelled AD shock is higher than the true one upon impact and more persistent thereafter, while the response of the interest rate to an AD shock is underestimated. In the low-right corner, notice that the estimated std of a MP shock, which is zero in the DGP, is a substantial 0.35. Therefore in the variance decomposition of all variables the role of MP shocks, which is truly zero, is estimated to be positive. The price puzzle has warning proportions. The response of output to a misspecified MP shock is highly persistent, reflecting the fact that the misspecified MP shocks are negatively correlated with the true technology shocks.

In the second experiment a stochastic element is added to the behavior of the monetary authorities by setting $\sigma_{MP} = 1$. The results are displayed in Figure 2. As previously argued, the misspecified system is no longer VAR(1) when

[9] The optimal value of the parameters of the Taylor function has a closed form solution given in Svensson (1997).

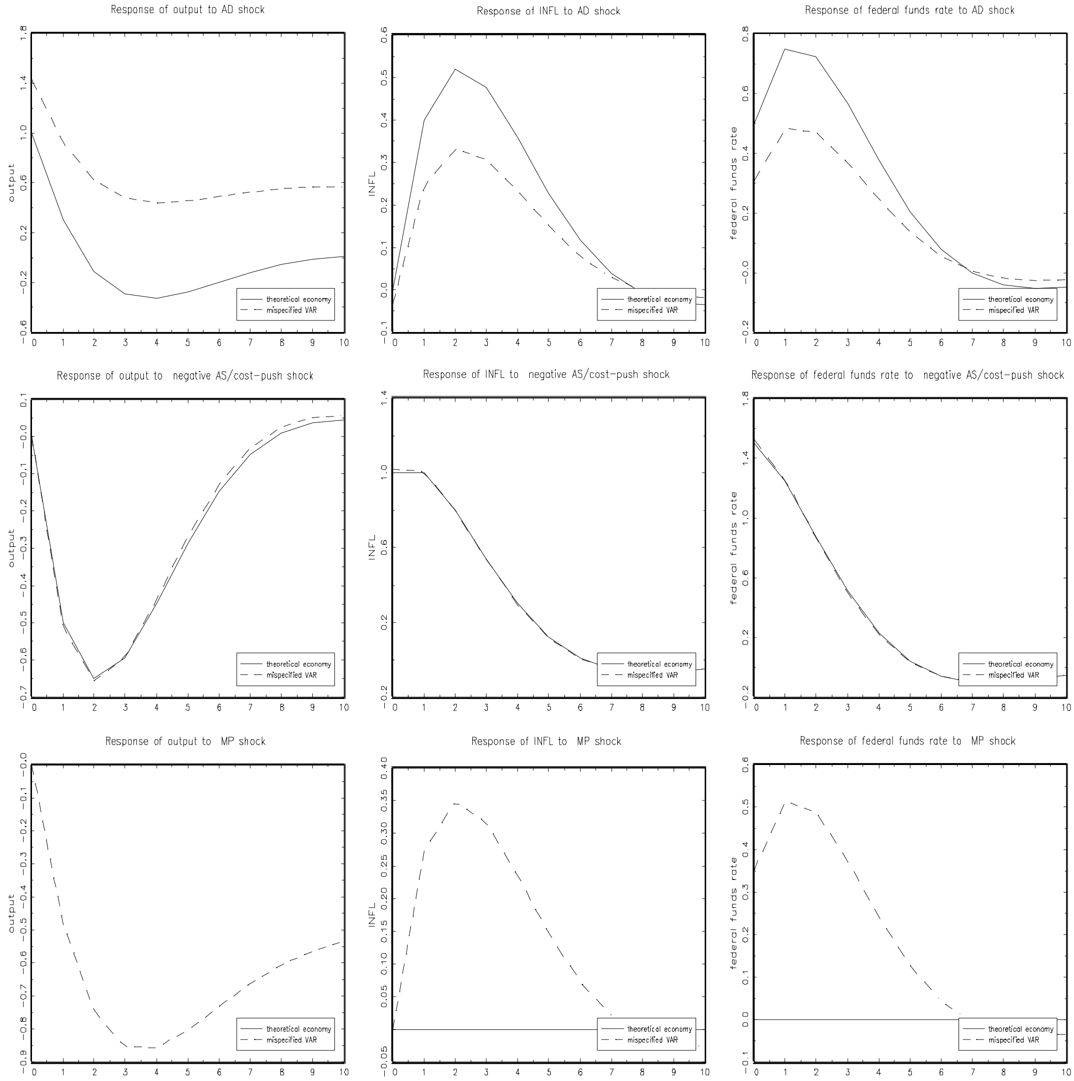

Figure 1: True impulse responses (thick line) and impulse responses from misspecified VAR (dashed line). $\sigma_{MP} = 0$.

$\sigma_{MP} > 0$, and impulse responses for the misspecified VAR have to be obtained numerically if we are to give the misspecified model its best chance.[10] Portmanteau test of residual correlation is first passed at four lags, so a VAR(4) is fit to the misspecified system. If less than four lags are chosen, the misspecifications maintain the same qualitative pattern but become larger. The response of the interest rate to an AD shock is underestimated, again as expected. Responses to MP shocks are once again those that display the most obvious misspecification. On the low-right corner, the std of the MP shock is overestimated. The price puzzle is substantial and can now be confronted with the true behavior of inflation in response to a MP shock. As for the response of output to a MP, not only is the size of the response overestimated, but the response is much longer lived than the true one (the reason being that the retrieved MP shocks are negatively correlated with the true technology shocks). These results extend without surprises to different combinations of ρ, σ_N, σ_{MP}.

In fact, the main results are valid if potential output is a non-constant deterministic function of time while the only exogenous variable in the misspecified VAR is a constant. In that case equation (10) does not change, while $\sigma^*_{MP} > 0$, since no combination of the included variables is perfectly correlated with the output gap: therefore there will be a price puzzle. Simulations using deterministic trends have shown that the effects of MP shocks in this case look much like in Figure 2.

2.3 Robustness of the main results to modifications in the model

The adoption of a completely specified model has allowed us to quantify the misspecifications. However, the main results (price puzzle, overestimation of

[10] The missspecified impulse responses are obtained by fitting a VAR on data generated by the DGP. The first 500 observations are not used in estimation. The VAR is estimated on 10000 observation to eliminate parameter uncertainty.

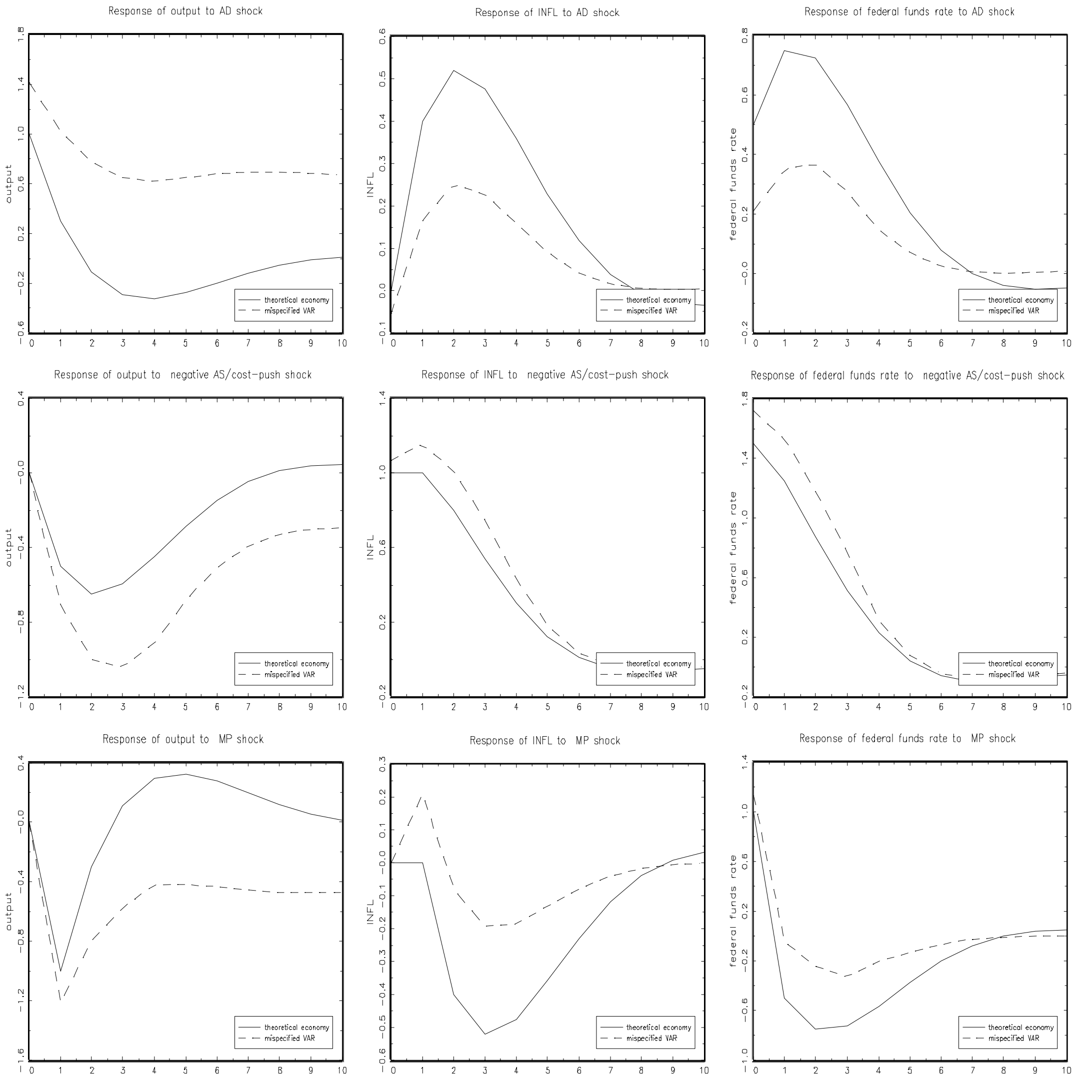

Figure 2: True impulse responses (solid line) and impulse responses from misspecified VAR(4) (dashed line). $\sigma_{MP} = 1$.

the variance of MP shocks) hold in a wide class of models. The models in this class have a reduced form solution characterized by[11] i) inflation responds with a lag and positively to the output gap ii) the monetary policy authority can affect inflation with no less than two lags iii) the output gap appears with a positive coefficient in the monetary policy function and y_t^g cannot be reduced to a linear combination of: a constant, Y_t, π_t, and varibles dated $t-1$ or earlier.

Overestimation of MP shocks follow from the fact that the output gap appears in the true policy function

$$i_{t+1} = \gamma_y y_{t+1}^g + ... \tag{19}$$

Since the output gap is omitted in the misspecified model, assumption iii) implies that the fit of the equation must deteriorate (notice that i and y^g are both dated $t+1$, so the forecasting power need not deteriorate). Therefore $\sigma_{MP}^* > 0$.

The price puzzle generates from the fact that the policy instrument appears in the misspecified inflation equation to pick up the role of the omitted output gap. If the reduced VAR form (only variable dated t or earlier in the right-hand-side) of the true inflation equation is

$$\pi_{t+1} = \alpha_y y_t^g + ..., \tag{20}$$

and monetary policy cannot affect inflation contemporaneously or at one lag, then, rearranging (19) and plugging it in (20), we'll find a price puzzle feature in the misspecified inflation equation, since $\frac{\partial \pi_{t+1}}{\partial i_t} = \frac{\alpha_y}{\gamma_y} > 0$ and $\sigma_{MP}^* > 0$.

Notice that these assumptions can be satisfied by models that

- Include microfounded and/or and forward looking relationship, as, for example, in Svensson (2000a and 2000b) and Clarida, Gali and Gertler (1999).

[11] The conditions are sufficient, not necessary.

- Have a fair amount of contemporaneous reactions. For example, technology shocks are allowed to affect all variables and potential output need not be exogenous. The only contemporaneous responses ruled out by assumptions are: inflation to AD shocks and all variables (except the interest rate) to MP shocks. In fact, the assumptions only guarantee that MP shocks can be identified using short-run restrictions, while all other shocks may not be.[12]

- Include a more complex loss function for the monetary authority, interest rate smoothing being a particularly interesting example.

The key assumption that MP shocks affect output with a lag and inflation with a longer lag is strongly supported in empirical work, including VAR studies, and is commonly incorporated in macro models for monetary policy analysis (for example the MPS macro model for the US).[13]

3 Solving the puzzle on US data

The strategy to test the hypothesis presented so far is straightforward:

1. Estimate a three variable VAR (the misspecified VAR) including: output (log of real GDP), CPI inflation, federal funds rate (same identification ordering).

2. Estimate the same VAR but with a measure of output gap rather than output. (I call this VARgap).

3. Compare the impulse responses of the two VARs and check whether they behave as predicted by the analysis of the Svensson model.

[12] Canova and Pina (1999) have an example of misspecification arising when the econometrician imposes short-run restrictions while the DGP does not have enough restrictions on contemporaneous responses to identify any shock.

[13] See, for example, Christiano, Eichembaum and Evans (1998) and Clarida, Gali and Gertler (1999).

In relation of the second point, a VAR including: a measure of output gap, inflation and federal funds rate is estimated in Rudebusch and Svensson (1999) and it does not produce a significant price puzzle. One of the contributions of this paper is to rationalize their finding.

As a measure of the output gap I use the series of capacity utilization built by the Federal Reserve Board.[14] Since capacity utilization is expressed as a percentage of full capacity for the manufacturing sector, a scale adjustment was used to account for the fact that industrial production (manufacturing) is more volatile than GDP. Therefore the series used in estimation is capacity$\times$0.5.[15] The response of a variables to a given shock in the two VARs are plotted on the same graph. A VAR(3) was estimated in all cases.[16] I checked the robustness of the results to different lag length structure (range 1-8) and starting sample dates (1960-1980). Switching capacity and inflation in the identification ordering also has no effect on either impulse responses or variance decompositions, as predicted by the Svensson model. Using the log of prices instead of inflation does not change any result in this and later sections. Figure 3 plots impulse responses for the two VARs estimated on the sample 1970:1 2000:2 (unless otherwise stated all figures are produced with a VAR(3) on the sample 1970:1 2000:2).

Variance decomposition for VARgap and for the misspecified VAR are presented in Figure 4 and Figure 5 respectively.

[14] Data description:
Capacity utilization is seasonally adjusted. It is available at FRED data base, http://www.stls.frb.org/fred/data/business/cumfg
The federal funds rate series is also taken from FRED, aggregated from monthly (averages), available at http://www.stls.frb.org/fred/data/business.
All other series are from the IMF database: GDP at constant prices (base year 1995), sa, CPI (all items), sa, were logged before estimation. All series used in this paper are available in an E-views workfile (please request them at nepgi@hhs.se).

[15] The number 0.5 is the result of the following computations. I assume that the output gap for GDP (in logs) is a multiple of the output gap in manufacturing. That is $y_t^g = \alpha y_t^{g,manufacturing}$, implying that $\alpha = std(y_t^g)/std(y_t^{g,manufacturing})$. Using data on sa industrial production and taking std of deviations from linear trends, I estimate $\alpha = 0.5$.

[16] The Schwarz and HQ criteria both select two lags for VARgap and, respectively, two and four lags for the misspecified VAR.

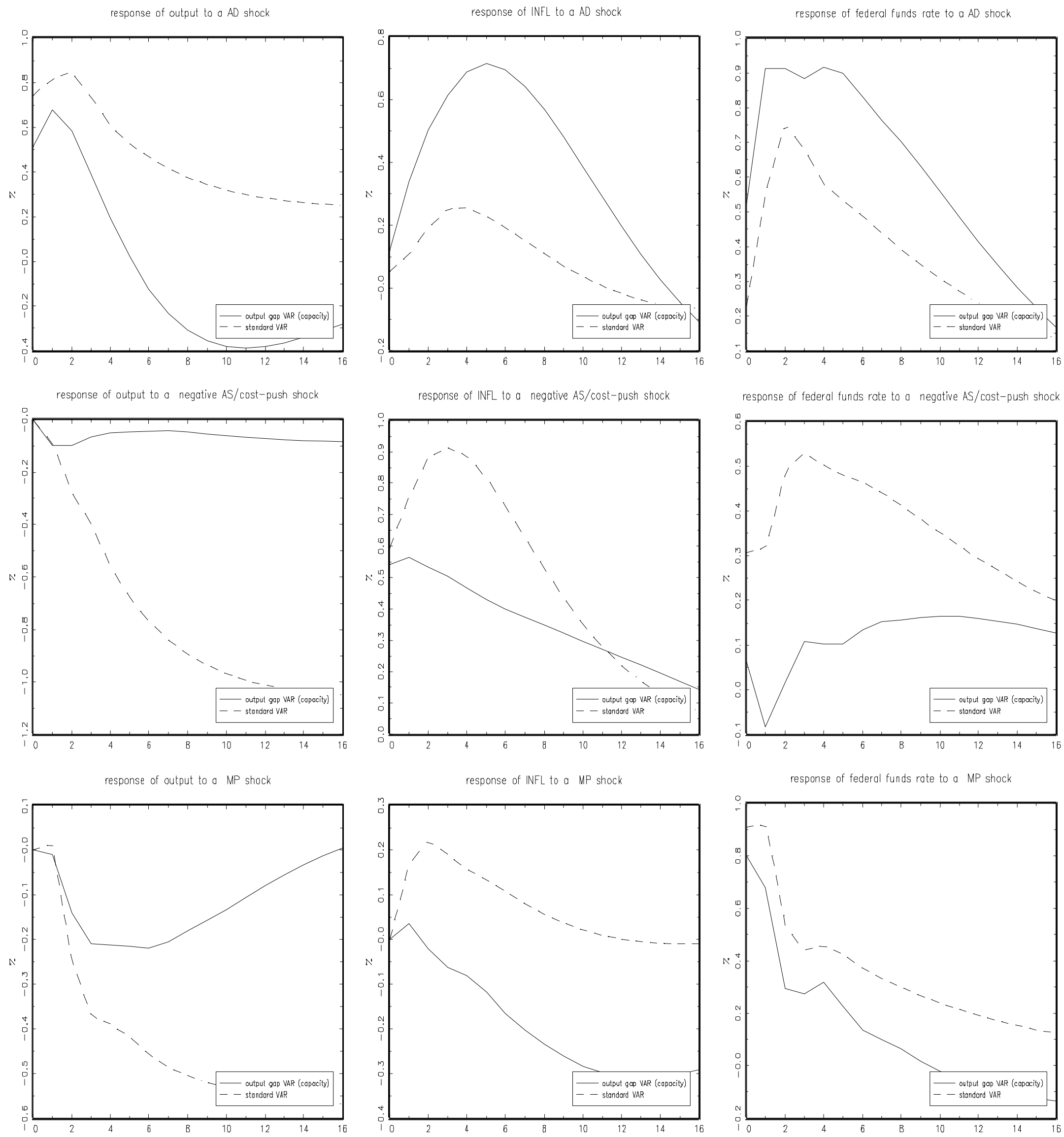

Figure 3: Impulse responses for misspecified VAR (real GDP, inflation and federal funds rate. Dashed line) compared with those from VARgap (output gap rather than output. Solid line). US data, quarterly, sample 1970:1 2000:2

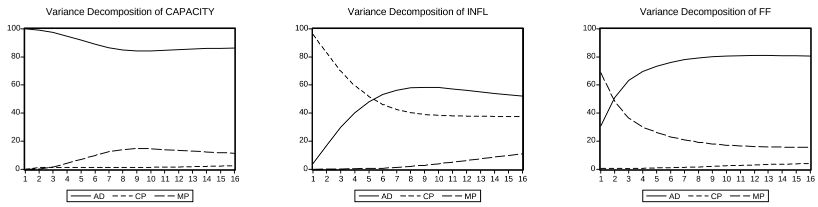

Figure 4: Variance decomposition for VARgap.

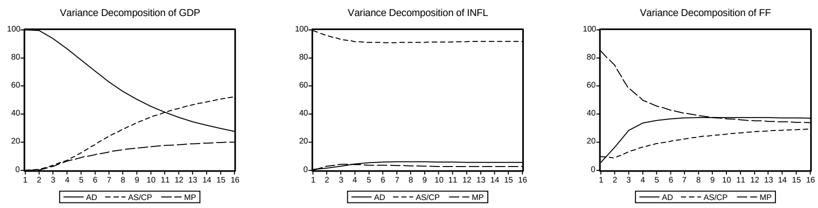

Figure 5: Variance decomposition for the misspecified VAR.

The reader is invited to compare Figures 3 and 2: in all nine cases the theoretical model correctly predicts whether the impulse response of VARgap lies above or below the impulse response of the misspecified VAR. I wish to underline the following results:

1. There is no price puzzle in VARgap, while there is a huge price puzzle in the misspecified VAR.
2. The response of the federal funds rate to an AD shock is higher in VARgap, even though the AD shock has a lower standard deviation.
3. The responses of output gap to all shocks are shorter-lived.
4. The std of MP shocks is 12% lower in VARgap, and standard deviations of the inflation equation and of the federal funds rate equation (in reduced form) are 7% and 11% lower. Therefore VARgap is expected to produce superior forecasts.

5. Monetary policy looks much more endogenous as the percentage of the federal fund rate forecast error variance due to MP shocks is substantially reduced in VARgap.

6. The share of MP shocks in the variance decomposition of output gap in VARgap is less (one half at the 16th lag) then in the decomposition of output in misspecified VAR.

7. The share of MP shocks in the variance decomposition of output in the misspecified VAR grows with the forecast horizon, as predicted (the reason being that the labelled MP shocks are correlated with technology shocks). In contrast, MP shocks in VARgap display no such behavior (the result doesn't change at forecast horizons longer than four years).

I have considered alternative measures of output gap, namely log deviations from a linear and from a quadratic trend and, for the fun of it, the cycle component of HP filtered log output.[17] All the main results are robust to the choice of the output gap proxy. However, using capacity utilization (and HP filter), nearly identical results are obtained for every reasonable choice of lags (range 1-8), while deviations from linear and quadratic trends (which remain highly persistent) produce a price puzzle for some choices of lags (four or higher). Capacity utilization produces the best fit in all equations.

3.1 Extending the VAR to include output: Technology shocks

The Svensson model admits both a three variable representation and four variable representations (see Section 2.1). So far three variable VARs have been estimated. Of course there is no reason why output should be excluded from

[17] The HP filter is two-sided, therefore the filtered data should not be used in regression analysis, since they will lead to inconsistent estimates.

the VAR as long as a measure of output gap is included. In this case technology shocks become part of the picture. The Svensson model then predicts that all responses to shocks other than technology shocks should be the same as in the three variable VARgap. Two orderings suggested by the model are: output gap, output, inflation, federal funds rate, and potential output, output, inflation, federal funds rate. I choose the second because it has the advantage of being correct even if technology shocks do affect the output gap (conditional, of course, on a good measure of potential output). It turns out that all results are robust to the choice of using output gap instead of potential output.[18] I construct potential output (logged) from real GDP and the measure of capacity used in the previous section.[19]

The impulse responses and variance decompositions are shown in Figure 6 and Figure 7. The following results stand out:

1. The inclusion of output does not produce any change worth noticing on any of the responses (the comparison is with the three variable VARgap). This is encouraging evidence that the shocks in VARgap had been identified correctly.

2. Impulse responses and variance decomposition indicate that natural output is an exogenous unit root process (AR modelling suggests that a random walk with drift is a good description).

3. The response of output to a technology shock is very similar to the response of natural output, while inflation and the federal funds rate have small and insignificant responses to technology shocks. That is, technology shocks do not affect the output gap, supporting the identification:

[18] The first ordering retrieves the shocks in the order: AD, technology, CP, MP. The second in the order: technology, AD, CP, MP.

[19] Log of natural output is defined as $y^n = y - (capacity/100)$, where capacity is obtained by multiplying the original series by 0.5, as motivated in Section 3.

output gap, output, inflation, federal funds rate. In fact, results from the two alternative identifications are fully compatible.

4. Variance decompositions also confirm the results obtained for the three variable VARgap. The share of technology shocks in the decomposition of inflation and federal funds rate is negligible, as predicted by the model. The share of technology shocks in the variance decomposition of output is substantial at all horizons (almost always above 50%).

5. Section two showed that, in the theoretical framework, the MP^* shocks retrieved from the misspecified VAR are positively correlated with the true AD shocks and negatively correlated with the true technology shocks. Using technology shocks and AD shocks taken from the four variable VAR and MP^* shocks from the misspecified VAR, this prediction can be tested and is in fact correct: $corr(\epsilon_{AD}, \epsilon^*_{MP}) = 0.35$, $corr(\epsilon_N, \epsilon^*_{MP}) = -0.32$.

6. The impulse responses at long lags (five years or longer, not shown) show that the long run effect of technology shocks on output is one to one, while all other shocks have a zero long run effect. Thus, although the identification is based on short-run restrictions, the results are fully consistent with the hypothesis that only technology shocks have a long-run impact on output.

4 Why does a commodity price index solve the price puzzle? Discriminating between two alternative explanations

This section argues that the commodity price index solves the price puzzle mainly because it contains useful information about the output gap, not because

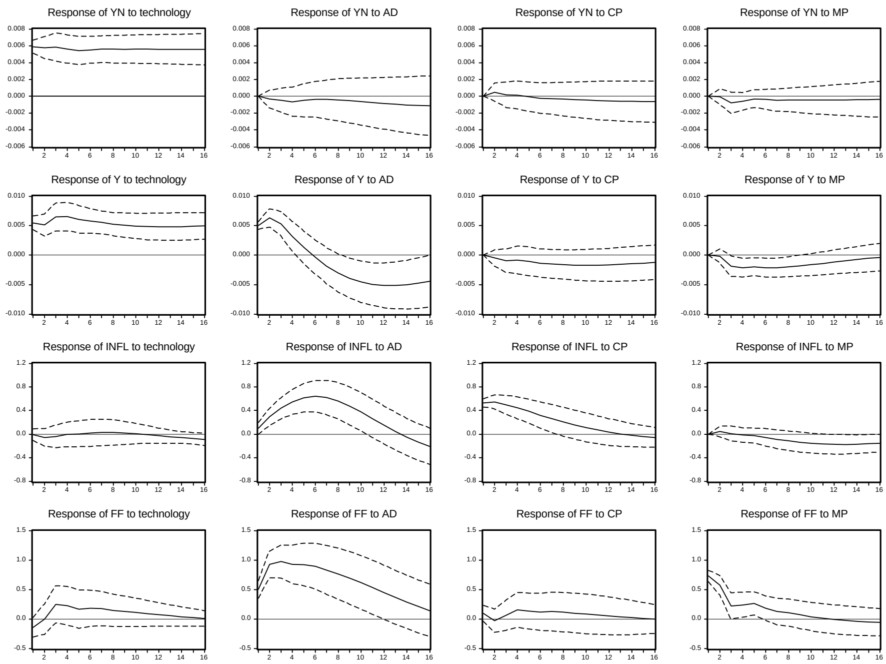

Figure 6: Impulse responses for the VAR: natural output, output, inflation, federal funds rate. Error bands include four standard deviations.

it is useful in forecasting inflation. PcomCEE[20] and capacity utilization do tend to move together (correlation 0.58 on the sample 1970:1-1998:4).[21]

The standard explanation implies that the price puzzle should disappear when a good leading indicator of inflation is included in the VAR. Since commodity prices have added value in predicting inflation, commodity prices should solve the price puzzle. The puzzle does almost completely disappear in a four variable VAR with the standard ordering: output, inflation, PcomCEE, federal funds rate.

But other powerful leading indicators of inflation should also go at least

[20] I call PcomCEE the index used by Christiano, Eichenbaum and Evans (1998).

[21] The publication of the commodity price index used in Christiano, Eichenbaum and Evans (1998) has been discontinued around 1996. The series was used by the Department of Commerce as a leading indicator. Data up to 1998Q4 have been kindly provided by Charles Evans, who constructed the last few data points following the procedure used by the DOC.

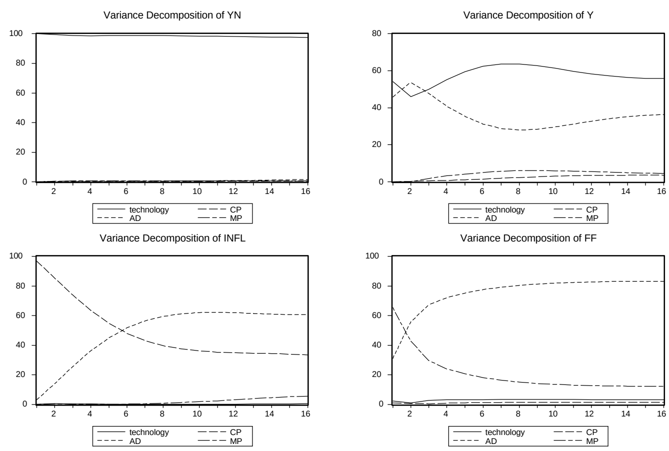

Figure 7: Variance decomposition for VAR4 (natural output, output, inflation, federal funds rate).

some way in solving the puzzle, if this theory is correct. For example, a long interest rate should react quickly to news of future inflation. This suggested using the yield on ten year government bonds instead of commodity prices. The results are surprising: the price puzzle is very large, as large as if this variable is omitted.

Next I estimate a four variable VAR including (in this order): output, inflation, inflation forecast, federal funds rate. Inflation forecast is the forecast of average inflation during the next four quarters, a reasonable time horizon for policy makers. The forecasts are produced with a VAR estimated recursively.[22] The resulting forecast series is then included in the four variable VAR in place of commodity price. The forecast do have some value, since in response to a forecast shock inflation grows monotonously. The response is significant at the 5% level for several quarters. Based on the standard explanation of the puzzle,

[22] The variables are: inflation, log of real output, commodity price index, federal funds rate and the three year yield on government bills. The VAR is first estimated on the sample 1960:1 to 1970:1 and the forecast is produced for average inflation during the periods 1970:2 to 1971:1. The VAR is then re-estimated adding one observation, another forecast is produced and so on.

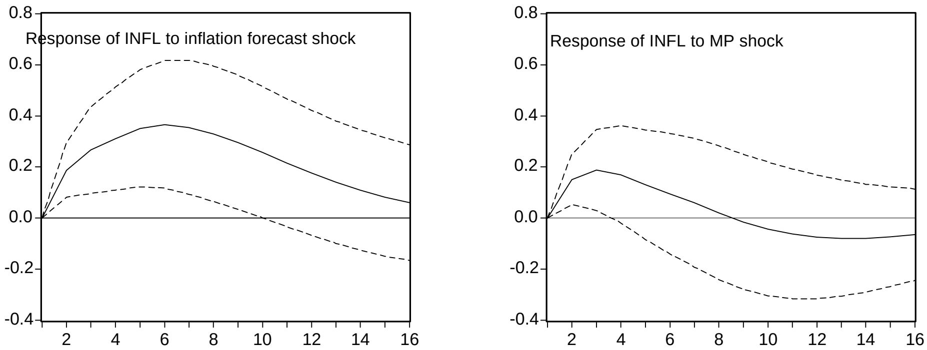

Figure 8: Response of inflation to news of future inflation and to a MP shock in a VAR including the forecast of future inflation from the ASA/NBER survey.

the inclusion of the inflation forecasts in the VAR should eliminate or at least mitigate the puzzle. In fact it doesn't help a bit. The response of prices to a MP shock (not reported) is unaffected. The same exercise is repeated with a forecast from a different source, the ASA/NBER Survey.[23] I take the mean across forecasters of the forecast of average inflation during the next four quarters (the same variable as before). The response of inflation to an expectation shock is highly significant (the error bands in Figure 8 are for ± 2 *std*), but the price puzzle remains substantial in both size and time extension.

These figures do not change if the forecast of future inflation is substituted with future inflation itself (one or two quarters ahead), with a little white noise error added to avoid having a singular variance-covariance matrix. Also, they do not change much using a standard "price of commodity" (as opposed to the leading indicator used by Christiano, Eichembaum and Evans (1998)), or the price of intermediate goods or industrial prices.

[23] Also known as the Survey of Professional Forecasters. Joutz and Stekler (2000) find that "The FED forecasts (*of GNP and GNP deflator*) were not significantly different from the predictions of ASA/NBER surveys". Since forecasts of CPI inflation are only available from 1981, I take the forecasts of GNP deflator inflation. Data and details are available at http://frb.libertynet.org/files/spf

In order to further test the claim, I test the following two hypothesis (the statements are for the null):

1. Once capacity is included in the Fed reaction function, PcomCEE is redundant.

2. Once PcomCEE in included in the Fed reaction function, capacity is redundant.

The testing procedures start with a model that nests both: federal funds rate regressed on a constant, three lags of itself, contemporaneous and lagged (three lags) values of inflation, output, PcomCEE, capacity (sample 1970:1 1998:4). The p-value for the F-statistic that commodity prices are redundant is 0.18. On the other hand, the hypothesis that capacity is redundant is clearly rejected (p-value 0.0003). In fact a better fit is obtained in the equation above excluding both output and PcomCEE than excluding capacity only.

Some researchers have found that the price puzzle almost completely disappears in some larger VARs that do not include a commodity price. I conjecture that this is due to some linear combination of the regressors being highly correlated with measures of output gap. For example, capital formation (logged and linearly detrended) has a correlation of 0.79 with linearly detrended log output (1970-2000). The correlation of unemployment rate [24] and capacity utilization is of 0.74. In fact, the unemployment rate used instead of capacity (ordered in any position, except the last) does a rather good job at solving the price puzzle, with inflation negative after four lags. Profits and all other strongly cyclical variables are likely candidates.

[24] Standardized unemployment rate, sa, OECD data base.

5 Conclusions

This paper argues that the finding of a positive response of inflation to a contractionary MP shock (price puzzle) in VARs designed for monetary policy analysis may not be due to monetary authorities having better forecasts than those produced by the VAR. Rather, it may be due to the omission of a measure of output gap in the VAR. This omission is shown to produce a price puzzle in a wide class of models that would display no such effect if correctly estimated and identified. The key requirement is that monetary policy affects output with a lag and inflation with a longer lag, a hypothesis strongly supported by empirical evidence. Using a model due to Svensson, it is shown that the omission of output gap also leads to overestimation of the variance of monetary policy shocks and to incorrect identification of the response of monetary policy to all the shocks in the economy. Moreover, the importance of monetary policy shocks in the variance decomposition of output is overestimated, and their effects are estimated to be longer than they actually are. The spurious appearance of a price puzzle is shown to be guaranteed in a wide class of models.

When the implications of the theoretical analysis are tested on US data all the main predictions are confirmed: the comparison of two three-variable VARs, one of which includes output rather than the output gap, gives the predicted results in terms of impulse responses and variance decompositions. In particular, MP shocks do not cause a price puzzle, have shorter-lived effects on output and less relevance in the variance decomposition of all variables. Thus very small VARs (three of four variables) can reproduce the results of much larger systems concerning the effects of MP shocks, while permitting identification of all shocks in the economy.

Potential output and technology shocks are also found to behave as assumed in the model. Technology shocks identified with short run restrictions taken from the model yield predictions consistent with the model and with the assumption (not imposed) that only technology shocks affect output in the long run. While the identification assumptions used in this paper are those suggested by a specific model and therefore need not be believable to everyone, they do allow a more informative test of this model since monetary policy shocks are not the only identified shocks.

Finally, it is argued that the effectiveness of a commodity price index in solving the puzzle does not depend on its being useful in predicting inflation. Rather, it is probably due to its fairly high correlation with the most popular measures of output gap. Using a measure of output gap (or potential output) is not only theoretically more appealing, but also leads to a better fit of the reaction function.

References

[1] Bagliano, F. C.; and Favero, C. A., 1998, "Measuring monetary policy with VAR models: An Evaluation," European Economic Review, 42, 1113-1140.

[2] Ball, L., 1999, "Policy Rules for Open Economies, " in John B. Taylor "Monetary Policy Rules", University of Chicago Press.

[3] Blanchard, O.; and Quah, D., 1989, "The Dynamic Effects of Aggregate Demand and Supply Disturbances, " American Economic Review, 79, 655-673.

[4] Canova, F.; and Pina, J. P., 1999, "Monetary Policy Misspecification in VAR models," working paper, Universitat Pompeu Fabra, Barcelona. Available at www.econ.upf.es

[5] Canova, F., 1998a, "Detrending and Business Cycle Facts," Journal of Monetary Economics, 41, 475,-512.

[6] Christiano, L.J., Eichenbaum M.; and Evans, C. L., 1998, "Monetary Policy Shocks: What Have We Learned and to What End?," NBER WP #6400.

[7] Clarida, R.; and Gali, J., 1994, "Sources of Real Exchange Fluctuations: How Important are nominal shocks?," Carnegie-Rochester Conference on Public Policy, 41, 1-56.

[8] Clarida, R.; Gali, J.; and Gertler, M., 1999, "The Science of Monetary Policy," Journal of Economic Literature 37, 1661-1707

[9] Favero, C., 2000, "Applied Macroeconometrics," Oxford University Press, Oxford.

[10] Gali, J., 1992, "How Well Does the IS-LM Model Fit Postwar US Data?," Quarterly Journal of Economics, 709-738.

[11] Leichter, J.; and Walsh, C., 1999 "Different Economies, Common Policies: Policy Trade-offs under the ECB," working paper. Available at http://econ.ucsc.edu/~walshc

[12] Lutkepohl, H., 1993, "Introduction to Multivariate Time Series Analysis", second edition, Springer-Verlag.

[13] Leeper, E. ; Sims C.; and Zha, T., 1996, "What Does Monetary Policy Do?," Brookings Papers on Economic Activity, vol. 2, 1-78.

[14] Joutz, F; and Stekler, H.O., 2000, "An Evaluation of the Predictions of the Federal Reserve", International Journal of Forecasting, 16, 17-38.

[15] Judd, J. P.; and Rudebush G. D., 1998, "Taylor's Rule and the Fed:1970-1997," Economic Review, 3, Federal Reserve Bank of San Francisco.

[16] Romer, D., 2000, "Keynesian Macroeconomics without the LM curve," Journal of Economic Perspectives, 14, 149-169.

[17] Rudebusch, G. D. ; and Svensson, L. E. O., 1999, "Policy Rules for Inflation Targeting, " in John B. Taylor "Monetary Policy Rules", University of Chicago Press.

[18] Sims, C., 1992, "Interpreting the macroeconomic time series facts: The effects of monetary policy," European Economic Review, 36, 2-16.

[19] Stock, J. H; and Watson, W., 1999, "Forecasting Inflation," Journal of Monetary Economics, 44, 293-336.

[20] Svensson, L. E.O., 1997, "Inflation forecast targeting: Implementing and monitoring inflation targets," European Economic Review, 41, 1111-1146.

[21] Svensson, L. E.O., 1999, "Inflation Targeting as a Monetary Plicy Rule," Journal of Monetary Economics, 43, 607-654.

[22] Svensson, L. E.O., 2000a, "Open-Economy Inflation Targeting," Journal of International Economics, 50, 155-183.

[23] Svensson, L. E.O., 2000b, "The Zero Bound in an Open Economy: a Foolprof Way of Escaping from a Liquidity Trap", NBER working paper n.7957.

6 Appendix: the correct and misspecified representation if the GDP has monetary policy shocks

The correct VAR.

$$\begin{bmatrix} y^g_{t+1} \\ \pi_{t+1} \\ i_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \gamma_y & \gamma_\pi & 1 \end{bmatrix} \begin{bmatrix} \beta_y & \beta_r & -\beta_r \\ \alpha_y & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y^g_t \\ \pi_t \\ i_t \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \gamma_y & \gamma_\pi & 1 \end{bmatrix} \begin{bmatrix} \epsilon^{AD}_{t+1} \\ \epsilon^{CP}_{t+1} \\ \epsilon^{MP}_{t+1} \end{bmatrix}$$

The misspecified VAR(1)

Use $y_t^g = Y_t - Y_t^N$ in the previous VAR, expand out the Y_{t+1}^N and Y_t^N terms and use $Y_{t+1}^N = \rho Y_t^N + \epsilon_{t+1}^N$.

$$\begin{bmatrix} Y_{t+1} \\ \pi_{t+1} \\ i_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \gamma_y & \gamma_\pi & 1 \end{bmatrix} \begin{bmatrix} \beta_y & \beta_r & -\beta_r \\ \alpha_y & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} Y_t \\ \pi_t \\ i_t \end{bmatrix} + \begin{bmatrix} -\beta_y + \rho \\ -\alpha_y \\ -\gamma_y \rho \end{bmatrix} Y_t^N +$$
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \gamma_y & \gamma_\pi & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{t+1}^{AD} + \epsilon_{t+1}^N \\ \epsilon_{t+1}^{CP} \\ \epsilon_{t+1}^{MP} - \gamma_y \epsilon_{t+1}^N \end{bmatrix}$$

Eliminate Y_t^N rearranging the Taylor rule

$$Y_{t+1}^N = Y_t - \frac{1}{\gamma_y} i_t. + \frac{\gamma_\pi}{\gamma_y}\pi_t + \frac{1}{\gamma_y}\epsilon_t^{MP},$$

to obtain the final representation

$$\begin{bmatrix} Y_{t+1} \\ \pi_{t+1} \\ i_{t+1} \end{bmatrix} = \begin{bmatrix} \rho & (\rho - \beta_y)\frac{\gamma_\pi}{\gamma_y} + \beta_r & -[(\rho - \beta_y)\frac{1}{\gamma_y} + \beta_r] \\ 0 & 1 - \alpha_y \frac{\gamma_\pi}{\gamma_y} & \frac{\alpha_y}{\gamma_y} \\ 0 & \beta_r \gamma_y - \beta_y \gamma_\pi + \gamma_\pi - \frac{\alpha_y}{\gamma_y}\gamma_\pi^2 & -\beta_r \gamma_y + \beta_y - \frac{\alpha_y}{\gamma_y}\gamma_\pi \end{bmatrix} \begin{bmatrix} Y_t \\ \pi_t \\ i_t \end{bmatrix} +$$
$$\begin{bmatrix} \epsilon_{t+1}^{AD} + \epsilon_{t+1}^N \\ \epsilon_{t+1}^{CP} \\ \gamma_y \epsilon_{t+1}^{AD} + \gamma_\pi \epsilon_{t+1}^{CP} + \epsilon_{t+1}^{MP} \end{bmatrix} + \begin{bmatrix} \rho - \beta_y \\ -\alpha_y \\ -\gamma_y \beta_y - \gamma_\pi \alpha_y \end{bmatrix} \epsilon_t^{MP}.$$

If $\sigma_{MP} > 0$, a VAR(1) will have autocorrelated errors. Therefore the econometrician is likely to select a longer lag lenght. It will soon be proved that the myspecified system has a VARMA(2,1) representation. Therefore the misspecified system will have inferior fit and forecasting efficiency than the correctly specified one, even if a VARMA(2,1) is estimated.[25]

Proposition 1 *The misspeficied system Y, π, i has a VARMA(2,1) representation.*

Proof. *Write the DGP as*

$$\begin{aligned} \begin{bmatrix} 1 - \rho L & 0 \\ A_0 - A_1 L & B_0 - B_1 L \end{bmatrix} \begin{bmatrix} Y_t^N \\ Z_t \end{bmatrix} &= \begin{bmatrix} \epsilon_t^N \\ \epsilon_t^Z \end{bmatrix}, \text{ or, in compact notation,} \qquad (21) \\ F(L) X_t &= \epsilon_t, \text{ where } Z_t' = \{Y_t\ , \pi_t,\ i_t\},\ \epsilon^Z = \{\epsilon_t^{AD},\ \epsilon_t^{CP},\ \epsilon_t^{MP}\}'. \end{aligned}$$

[25] See Lutkepohl (1993), pag. 234.

Where L denotes the lag operator and

$$A_0 = \begin{bmatrix} -1 \\ 0 \\ \gamma_y \end{bmatrix},\ A_1 = \begin{bmatrix} \beta_y \\ \alpha_y \\ 0 \end{bmatrix},\ B_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\gamma_y & -\gamma_\pi & 1 \end{bmatrix},\ B_1 = \begin{bmatrix} \beta_y & \beta_r & -\beta_r \\ \alpha_y & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Premultiply both sides of (21) *by* $F(L)^{-1}$. *The resulting system for* Z_t *is*

$$Z_t = (B_0 - B_1 L)^{-1}(A_0 - A_1 L)(1 - \rho L)^{-1}\epsilon_t^N + (B_0 - B_1 L)^{-1}\epsilon_t^Z.$$

Premultiply both sides by $(B_0 - B_1 L)(1 - \rho L)$ *and rearrange*

$$Z_t = (B_0^{-1}B_1 + \rho I)Z_{t-1} - \rho B_0^{-1}B_1 Z_{t-2} - (A_0 - A_1 L)\epsilon_t^N + (1 - \rho L)B_0^{-1}\epsilon_t^Z,$$

where $-(A_0 - A_1 L)\epsilon_t^N + (1 - \rho L)B_0^{-1}\epsilon_t^Z$ *has a multivariate* $MA(1)$ *represenation of the form* $-(A_0 - A_1 L)\epsilon_t^N + (1 - \rho L)B_0^{-1}\epsilon_t^Z = (I + ML)u_t$, u *multivariate (three dimensional) white noise (Lutkepohl (1993), pag. 231).*

■