Measuring the Non-Linear Effects of Monetary Policy*

Regis Barnichon
CREI, Universitat Pompeu Fabra, CEPR

Christian Matthes
Federal Reserve Bank of Richmond

July 2014
(first draft: March 2014)

Abstract

While there exists a broad consensus about the average effect of monetary policy on output, there is still considerable uncertainty about possible asymmetric or non-linear effects of monetary policy. We propose a new method to identify the (possibly non-linear) effects of policy shocks by using Gaussian basis functions to parametrize the impulse response functions. We find that the effect of a monetary shock depends strongly on (i) the sign of the shock, (ii) the size of the shock, and (iii) the state of the business cycle at the time of the shock. A contractionary policy has a strong adverse effect on output, much stronger than linear estimates suggest, but an expansionary policy has, on average, no significant effect on output. An expansionary policy can have some expansionary effect on output, but only if the intervention is large and during a recession. Even so, a contractionary policy is always more potent than its expansionary counterpart.

JEL classifications: E24, E32

*We would like to thank Christian Brownlees, Davide Debortoli, Jordi Gali, Yuriy Gorodnichenko, Kris Nimark, Ricardo Reis, Barbara Rossi, Mark Watson, Yanos Zylberberg and seminar participants for helpful comments. We thank Lorenz Kueng for sharing the extended Romer and Romer monetary shock series. Barnichon acknowledges financial support from the Spanish Ministerio de Economia y Competitividad (grant ECO2011-23188), the Generalitat de Catalunya (grant 2009SGR1157) and the Barcelona GSE Research Network. The views expressed here do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System. Any errors are our own.

1 Introduction

There now exists a broad consensus on the average effect of monetary policy on output, and it is widely accepted that a monetary contraction (expansion) leads to a persistent decline (increase) in output. Surprisingly however, there is still little agreement about possible asymmetric or non-linear effects of monetary policy.

In fact, three questions at the core of monetary policy making are still unsettled. First, does monetary policy have asymmetric effects on economic activity? As captured by the string metaphor, a contractionary monetary shock is generally thought to have a much stronger effect –being akin to pulling on a string– than an expansionary shock –being akin to pushing on a string–.[1] However, empirical evidence has been inconclusive, and the asymmetric effect of monetary policy remains an open question. Second, does the effect of monetary policy vary with the size of the policy intervention? While it was theoretically suggested that big monetary shocks have smaller effects on real activity than small shocks,[2] there is again no conclusive evidence. Third, does the effect of monetary policy vary with the state of the business cycle? For instance, does the central bank have more room to stimulate output (without raising inflation) during recessions?

Providing uncontroversial answers to these questions has been difficult in part for one important technical reason: the standard approach to identifying the effect of monetary shocks –Structural Vector-Autoregressions (SVARs) with a recursive identification scheme–,[3] is linear and does not allow for asymmetry or non-linearity in the effect of monetary shocks.

To overcome this limitation, this paper proposes a new method to identify the (possibly non-linear) effects of monetary policy. In similar spirit to the structural VAR approach, we identify exogenous policy shocks from a recursive identification scheme. However, different from the SVAR approach, we allow the impulse responses to shocks to depend on the sign or the value of the identified structural shocks. Since this generalized model does not admit a VAR representation, we instead directly estimate the impulse response functions from the data by parameterizing the impulse response functions with Gaussian basis functions.

Our approach builds on two premises: (i) any mean-reverting impulse response function can be approximated to any degree of accuracy by a sum of Gaussian basis functions, and

[1] For instance, increasing borrowing costs may reliably slow an expansion, but that cheap capital need not stimulate investment in a downturn. Another reason could be the existence of downward price or wage rigidity. See Morgan (1993) for a discussion of possible reasons for asymmetry.

[2] This could be the case in models with menu costs (Ball and Romer, 1990), in which firms find optimal to adjust prices to big monetary shocks (making large monetary shocks neutral) but not to adjust to small monetary shocks (making small shocks non-neutral).

[3] See e.g., Christiano, Eichenbaum, and Evans (1999). The recursive identification scheme typically states that monetary policy does not affect the real economy instantaneously.

(ii), in practice, only a very small number of Gaussian functions are needed to approximate a typical impulse response function. Intuitively, the impulse response functions of stationary variables are often found to be monotonic or hump-shaped (e.g., Christiano, Eichenbaum, and Evans, 1999 or Uhlig, 2005). In such cases, a single Gaussian function can already provide an excellent approximation of the impulse response function.

Thanks to the small number of free parameters allowed by a Gaussian parametrization, it is possible to directly estimate the impulse response functions from the data, by maximum likelihood or Bayesian methods.[4] The parsimony of the approach in turn allows us to estimate more general non-linear models.

Consistent with the string metaphor, our findings point towards the existence of very strong asymmetry in the effect of monetary policy: a contractionary monetary policy shock generates a large and significant reduction in output, but an expansionary shock has a small and insignificant effect. As a result, conventional (linear) estimates may substantially over-estimate the real effect of expansionary policy and under-estimate the real effect of contractionary monetary policy.[5]

When we allow the effects of monetary policy to depend on the size of the intervention as well as the state of the business cycle, we find strong evidence in favor of both "size dependence" and "state dependence". While an expansionary policy has, on average, no significant effect on output, the effect varies markedly with the state of the cycle. An expansionary policy during a strong growth episode has no effect on output and only generates inflation. In contrast, an expansionary policy can have a stimulating effect on economic activity (at no inflationary costs) when the economy is in a recession. Moreover, the positive effect of an expansionary policy on output increases slowly with the size of the intervention, so that an expansionary policy is most effective if the intervention (i) takes place during a recession and (ii) is relatively large. However, an expansionary policy is always considerably less potent than its contractionary counterpart. In fact, the adverse effect of a contractionary policy on output is large and relatively independent of the state of the business cycle.

While the SVAR approach has allowed the emergence of a broad consensus about the average effect of monetary policy on output, with a monetary contraction leading to a persistent decline in output, there is little agreement about the asymmetric effect of monetary policy. For instance, while Cover (1992) finds evidence of asymmetric effects, Ravn and Sola (1996, 2004)

[4] Another important advantage of using Gaussian basis functions is that, in the context of Bayesian estimations, prior elicitation is much easier than with Bayesian estimations of standard VARs, because the coefficients to be estimated are directly interpretable as features of impulse responses.

[5] This latter observation is of particular relevance for the current debate on the appropriate timing for the lift-off of the fed funds rate, currently set at the zero lower bound. Our estimates suggest that an inappropriate (i.e., too strong or too early) increase in the fed funds rate could be a lot more costly (in terms of lost output) than conventional estimates suggest.

and Weise (1999) instead find nearly symmetric effects. This lack of consensus is perhaps not surprising: without a proper way to identify exogenous policy changes in a non-linear framework (as structural VARs allow in the linear case), identification of the asymmetric (or non-linear) effect of monetary shocks is difficult.

One strand of the literature has relied on threshold models to capture the stance of the business cycle or the state of monetary policy. However, such threshold models suffer two important drawbacks: (i) the threshold applies to a "switching variable", either observable, e.g., past unemployment or past output growth, or not,[6] which is not a structural shock but instead the function of past reduced-form shocks, and (ii) there is no "right" switching variable, with results sensitive to the particular choice of switching variable (Weise, 1999), making the interpretation of results difficult. In contrast, our approach identifies the monetary shocks and then estimates whether the economy reacts differently depending on the sign or value of the contemporaneous monetary shock itself (in this sense, the right switching variable is the contemporaneous structural monetary shock).[7] Our model with "state dependence" bridges with the threshold model literature, in that we allow the effect of policy to depend on the state of the business cycle, which is measure using an indicator variable similar to the threshold variable.[8] By being more general and by explicitly separating the roles played by sign-, size-, or state-dependence, our approach can help better identify the contribution of each different non-linearity.

Another approach, pursued by Cover (1992), DeLong and Summers (1988) and Morgan (1993), uses estimates of unanticipated money supply changes (as in Barro, 1977) to test whether the economy responds differently to the sign of the shock. However, unlike the modern VAR approach to identifying exogenous monetary shocks (and unlike our methodology), this approach does not allow the monetary policy rule to respond to the state of the economy, making the identification of shocks difficult.

Recently, Angrist et al. (2013) propose a novel semi-parametric time series estimator to evaluate the effect of monetary policy interventions and find asymmetric effects of monetary shocks. Despite very different approaches, both results tell a similar story with contractionary policy interventions being a lot more effective than expansionary ones.

Our use of Gaussian basis functions to approximate (and parametrize) impulse response

[6]See, among others, Beaudry and Koop (1993), Thoma (1994), Potter (1995), Kandil (1995), Koop, Pesaran and Potter (1996), Koop and Potter, (1998), Ravn and Sola (1996, 2004), Weise (1999).

[7]In threshold models, the coefficients are functions of *lagged* observable or *lagged* unobservable variables. In contrast, our approach isolates the *contemporaneous* effects of the policy intervention: a positive shock can have a different contemporaneous response coefficient (i.e., effect) than a negative shock.

[8]Relatedly, Tenreyro and Thwaites (2013) recently estimate the extent of "state dependence" by using Romer and Romer (2004) monetary policy shocks and building on Auerbach and Gorodnichenko (2012) threshold model. They find that monetary policy effectiveness declines in recessions.

functions relates to a large literature (mostly outside of economics) that relies on radial basis functions (of which Gaussian functions are one example) to approximate arbitrary multivariate functions (e.g., Buhmann, 2003) or to approximate arbitrary distributions using a mixture of Gaussian distributions (Alspach and Sorenson 1971, 1972, McLachlan and Peel, 2000). In economics, our parametrization of impulse responses relates to an older literature on distributed lag models and in particular on the Almon (1965) lag specification, in which the successive weights are given by a polynomial function. Although we rely on radial basis functions rather than polynomials, the work of Jorgenson (1966) suggests that ratios of polynomials could also be used to parametrize impulse response functions.[9]

Section 2 presents the empirical model, Section 3 presents the theoretical background behind our use of Gaussian basis functions; Section 4 describes possible parameterizations with Gaussian functions; Section 5 describes the estimation routine; Section 6, 7 and 8 present evidence for the existence of, respectively, asymmetry, size dependence and state dependence; Section 9 concludes.

2 Empirical model

Our starting point is a general description of the economy, in which the behavior of a system of macroeconomic variables is dictated by its response to past and present shocks. Specifically, denoting $\mathbf{Y}_t$ a vector of non-policy macroeconomic variables (e.g., output growth, inflation) and p_t a scalar variable that summarizes the policy stance, the economy is described by

$$\mathbf{Z}_t = \begin{pmatrix} \mathbf{Y}_t \\ p_t \end{pmatrix} \tag{1}$$

$$= \sum_{k=0}^{\infty} \mathbf{\Psi}_k(\varepsilon_{t-k})\varepsilon_{t-k} \text{ with } \varepsilon_{t-k} = \begin{pmatrix} \varepsilon_{t-k}^y \\ \varepsilon_{t-k}^p \end{pmatrix} \tag{2}$$

where boldface letters are used to indicate vectors or matrices of variables or coefficients. ε_t is the vector of structural innovations (and ε_t^p denote a monetary policy innovation) with $E\varepsilon_t = \mathbf{0}$ and $E\varepsilon_t\varepsilon_t' = \mathbf{I}$, and $\mathbf{\Psi}_k(\varepsilon_{t-k})$ is the matrix of lag coefficients.

Importantly, the model is not a priori linear and the coefficients of $\mathbf{\Psi}_k$ can depend on the value of the structural innovations ε_{t-k}. For instance, if $\mathbf{\Psi}_k$ depends on the sign of ε_{t-k}, the model will allow for asymmetric effect of structural innovations. In other words, (2) can be seen as a non-linear vector moving average representation of the economy.

[9] Jorgenson showed that a a ratio of polynomials, called a "rational distributed lag model", can approximate any distributed lag function (i.e., any impulse response function) to any degree of accuracy with fewer parameters than a finite distributed lag function. We preferred using Gaussian basis functions, since they seemed a more natural and more parsimonious way to approximate mean-reverting impulse response functions.

To ensure identification of the monetary policy innovations, in this paper, we assume that the standard recursive identification restriction is verified:[10] Monetary policy shocks are assumed to not affect macro variables within the current period, which implies that $\mathbf{\Psi}_0$'s last column is filled with 0 except for the last entry.[11]

If (2) admitted a VAR representation, the model could be estimated by first estimating a VAR $\binom{\mathbf{Y}_t}{p_t}$ and then inverting it to get the moving average representation (2). However, this standard approach is only possible if (2) admits a VAR representation, which imposes that the model is linear, i.e., $\mathbf{\Psi}_k(\varepsilon_{t-k}) = \mathbf{\Psi}_k$,[12] and prevents the study of possible non-linear effects of monetary policy.

3 Parameterizing impulse response functions using Gaussian basis functions

In this section, we propose an approach to estimate a non-linear model like (2). Rather than looking for a VAR representation of the dynamic system (2), our aim is to directly estimate (2), the "MA representation" of the economy. Because the number of free parameters $\mathbf{\Psi}_k$ in (2) is infinite, our strategy consists in parameterizing the impulse response functions, i.e., in positing a functional form for each impulse response function $\psi_{ij}(k)$ (with $\psi_{ij}(k)$ the row-i column-j coefficient of $\mathbf{\Psi}_k$, i.e., the impulse response function of variable i to shock j). To parametrize the impulse response functions, we will use Gaussian basis functions.

We first present the theoretical background supporting our parametrization, discuss the intuition and motivation behind our approach in practice, and finally present the different specifications that we will estimate.

3.1 Theoretical background

Our parametrization of the impulse response functions builds on the following theorem that states that any mean-reverting impulse response function on a compact interval can approximated with a sum of Gaussian functions.

Theorem 1: Let $f(x)$ be a continuous function over $[0, \infty[$ that satisfies $\int_0^\infty |f(x)|\, dx < \infty$.

[10] In the context of monetary shocks, this identification scheme is standard in the literature. See, among others, Bernanke and Blinder (1992), Bernanke and Mihov (1998), Christiano et al. (1999, 2005), Boivin and Giannoni (2006), and Rotemberg and Woodford (1997).

[11] As an alternative, we will also estimate our main specifications using the Romer and Romer (2004) shocks identified with a narrative approach and extended until 2008 by Coibion.

[12] In addition, the existence of a VAR representation imposes some restrictions on the $\mathbf{\Psi}_k$ coefficients. The matrices $\mathbf{\Sigma}_k \equiv \mathbf{\Psi}_0^{-1}\mathbf{\Psi}_k$ must be such that there exists $p > 0$ and matrices $\Phi_1, ..., \Phi_p$ such that $\mathbf{\Sigma}_k$ can be written as $\mathbf{\Sigma}_k = \Phi_k\mathbf{\Sigma}_{k-1} + ... + \Phi_p\mathbf{\Sigma}_{k-p}$, $\forall k > 0$.

Then, there exists a function $f_N(x)$ defined by

$$f_N(x) = \sum_{n=1}^{N} a_n e^{-(\frac{x-b_n}{c_n})^2}$$

with a_n, b_n, $c_n \in \mathbb{R}$ for $n \in \mathbb{N}$, such that $f_N(x) \underset{N\to\infty}{\longrightarrow} f(x)$ on every interior subinterval of $[0, \infty[$.

Proof: See Appendix.

Using Theorem 1, we can then approximate any mean-reverting (in the sense that $\int_0^\infty |\psi(k)|\, dk < \infty$) impulse response function $\psi(k)$ by a sum of Gaussian functions:

$$\psi(k) \simeq \sum_{n=1}^{N} a_n e^{-(\frac{k-b_n}{c_n})^2}$$

A limitation however is that we need the impulse response functions to be integrable, which restricts our approach to stationary series (so that $\mathbf{\Psi}_k \underset{k\to\infty}{\longrightarrow} \mathbf{0}$).[13,14]

3.2 Intuition and Motivation

The interest of our approach, and its use for studying the (possibly non-linear) effects of policy, rests on the fact that, in practice, only a very small number of Gaussian basis functions are needed to approximate a typical impulse response function.

Intuitively, impulse response functions of stationary variables are often found to be monotonic or hump-shaped (e.g., Christiano, Eichenbaum, and Evans, 1999 or Uhlig, 2005).[15] And in such cases, a single Gaussian function can already provide a good approximate description of the impulse response. To illustrate this observation, Figure 1 plots the impulse response functions of output, the price level and the fed funds rate to a monetary shock estimated from a standard VAR specification,[16] along with the corresponding Gaussian approximations with only *one* Gaussian function, i.e., using the approximation

$$\psi(k) \simeq a e^{-\frac{(k-b)^2}{c^2}}. \tag{3}$$

[13] Another (practical this time) limitation is that since we are estimating an *infinite* moving average model with a *finite* number of observations, the impulse responses must satisfy $\mathbf{\Psi}_k \underset{k\to\infty}{\longrightarrow} \mathbf{0}$).

[14] With non-stationary series, an option is thus to take a stationary transformation of the data (e.g., taking first-differences) so as to only include stationary series in the model.

[15] This is also the case in theoretical models, e.g., New-Keynesian models, in which the impulse response functions are generally monotonic or hump-shaped (see e.g., Walsh 2010).

[16] See Section 6 for the exact specification of the SVAR behind Figure 1. The VAR is specified with (log) industrial production, inflation and the fed funds rate. The impulse response for the price level is calculated from the response of inflation.

We can see that the one-Gaussian parametrization already does an excellent job at approximating the impulse responses implied by the VAR.

The small number of free parameters (only three per impulse response function; a, b and c), will have two important advantages. First, it will allow us to directly estimate the impulse response functions from the MA representation (2).[17] Second, it will allow us to introduce (and estimate) asymmetric or non-linear effects of monetary policy.

Another advantage of the one-Gaussian parametrization is the a,b and c coefficients can be very easily interpreted in light of the impulse-response function. In fact, with a one-Gaussian parametrization, the impulse response function is summarized by three parameters –peak effect, timing of peak effect, and persistence of peak effect–, which are generally considered the most relevant characteristics of an impulse response function.[18] As illustrated in Figure 2, parameter a is the height of the impulse-response, i.e., the maximum effect that a given shock has on the variable of interest, parameter b is the timing of this maximum effect, and finally parameter c captures how persistent the effect of the shock is.[19] Specifically, for the latter term c, it is easy to show that the amount of time τ required for the effect of a shock to be 50% of its maximum value is given by

$$\tau = c\sqrt{\ln 2.}$$

A final important advantage of (3) is that, in the context of Bayesian estimation, the ease of interpretation of the a,b and c parameters makes prior elicitation easier than in standard VARs, in which the VAR coefficients do not have such a direct economic interpretation.

4 Examples of parametrization with Gaussian functions

In this section, we consider different parametrization of interest, starting with linear models, using either a one-Gaussian or two-Gaussian approximation of the impulse responses, and then allowing for non-linearities with sign-dependence (asymmetry), size-dependence and state-dependence.

[17] For instance, with 3 variables, we only have one $9 * 3 = 27$ parameters to estimate to capture the whole set of impulse response functions $\{\mathbf{\Psi}_k\}_{k=1}^{\infty}$.

[18] For instance, when comparing the effects of monetary shocks across different specifications, Coibion (2012) focuses on the peak effect of the monetary shock, which in our framework is simply parameter a.

[19] Since parameter c_{ij} divides k, it can be seen as dilating or expanding time. The larger c_{ij}, the longer a given effect will persist.

4.1 Linear effects of policy shocks: one-Gaussian approximation

First, as a benchmark case, consider the general model (2) in the linear case, i.e., when the economy is described with

$$\mathbf{Z}_t = \sum_{k=0}^{\infty} \mathbf{\Psi}_k \varepsilon_{t-k}. \tag{4}$$

Using a one-Gaussian approximation of the impulse responses, the k lag coefficient $\psi_{ij}(k)$ (with $\psi_{ij}(k)$ the row-i column-j coefficient of $\mathbf{\Psi}_k$) is given by

$$\psi_{ij}(k) = a_{ij} e^{-\frac{(k-b_{ij})^2}{c_{ij}^2}}, \ \forall k > 0 \tag{5}$$

with a_{ij}, b_{ij}, c_{ij} some constants to be estimated.

In practice, to allow for more flexibility, we will let (5) only hold for $k > 0$, and we will leave the contemporaneous coefficient matrix $\mathbf{\Psi}_0$ unconstrained except that the elements above the main diagonal are zero (the structural identifying restriction) and that the diagonal elements are positive (a normalization).

4.2 Linear effects of policy shocks: two-Gaussian approximation

To allow for a richer set of impulse-responses, in particular to allow for an overshooting (or oscillating) pattern in the impulse-response, we can consider approximating the impulse response functions with two Gaussian functions, namely

$$\psi_{ij}(k) = a_{ij,1} e^{-\frac{(k-b_{ij,1})^2}{c_{ij,1}^2}} + a_{ij,2} e^{-\frac{(k-b_{ij,2})^2}{c_{ij,2}^2}}, \ \forall k > 0 \tag{6}$$

with $a_{ij,1}$, $b_{ij,1}$, $c_{ij,1}$ and $a_{ij,2}$, $b_{ij,2}$, $c_{ij,2}$ some constants to be estimated. Again, the contemporaneous coefficient matrix $\mathbf{\Psi}_0$ is left unconstrained.

4.3 Asymmetric effects of policy shocks

With a parametric representation of the impulse response functions in hand, it is simple to extend the approach to allow for asymmetries by allowing $\mathbf{\Psi}_k$ to depend on the sign of the structural shock and to take two possible values $\mathbf{\Psi}_k^+$ or $\mathbf{\Psi}_k^-$. Specifically, a model that allows for asymmetric effect of monetary shocks would simply write

$$\mathbf{Z}_t = \sum_{k=0}^{\infty} \mathbf{\Psi}_k^+ \mathbf{1}_{\varepsilon \geq \mathbf{0}} \varepsilon_{t-k} + \sum_{k=0}^{\infty} \mathbf{\Psi}_k^- \mathbf{1}_{\varepsilon < \mathbf{0}} \varepsilon_{t-k} \tag{7}$$

with $\mathbf{\Psi}_k^+$ and $\mathbf{\Psi}_k^-$ the lag matrices of coefficients for, respectively, positive and negative shocks. For instance, in a bivariate case, we would have

$$\mathbf{\Psi}_k^+ = \begin{pmatrix} \psi_{11}^+(k) & \psi_{12}^+(k) \\ \psi_{21}^+(k) & \psi_{22}^+(k) \end{pmatrix} \tag{8}$$

the coefficient matrix for positive shocks. A similar expression holds for $\mathbf{\Psi}_k^-$. We then have $\psi_{ij}(k) = \psi_{ij}^+(k)\mathbf{1}_{\varepsilon_i \geq 0} + \psi_{ij}^-(k)\mathbf{1}_{\varepsilon_i < 0}$.

Using a one-Gaussian approximation, the impulse-response function would be parametrized as

$$\psi_{ij}^+(k) = a_{ij}^+ e^{-\left(\frac{k-b_{ij}^+}{c_{ij}^+}\right)^2}, \ \forall k > 0 \tag{9}$$

with a_{ij}^+, b_{ij}^+, c_{ij}^+ some constants to be estimated. A similar expression would hold for $\psi_{ij}^-(k)$.

4.4 Size-dependent effects of policy shocks

With a parametric representation of the impulse response functions, it is also simple to allow for non-linearities (other than asymmetry). For instance, we can allow for size dependence, i.e., allow the marginal effect of a monetary shock to depend on the magnitude of the shock. Specifically, with a one-Gaussian approximation, one could estimate a model of the form

$$\begin{cases} \psi_{ij}(k) = a_{ij}(\varepsilon_{t-k}) e^{-\left(\frac{k-b_{ij}}{c_{ij}}\right)^2}, \ \forall k > 0 \\ a_{ij}(\varepsilon_{t-k}) = \alpha_{ij} + \beta_{ij} \left|\varepsilon_{t-k}\right| \end{cases} \tag{10}$$

with α_{ij}, β_{ij}, b_{ij} and c_{ij} parameters to be estimated. In this model, the "amplitude" of the impulse response (the parameter a_{ij}), i.e., the maximum effect of a shock, is a linear function of the size of the shock. Such a specification allows us to test whether the effect of monetary policy vary with the size of the policy intervention. Naturally, other (or more general) models are possible.[20]

4.5 Asymmetry and size-dependence

One can then combine the two previous specifications to construct a model that allows for both asymmetry and size-dependence. Specifically, using a one-Gaussian approximation, the

[20] For instance, one could also allow the mode of the impulse-response (the parameter b_{ij}) to depend on the value of the shock.

impulse response function following a positive innovation could be parametrized as

$$\begin{cases} \psi_{ij}^{+}(k) = a_{ij}^{+}(\varepsilon_{t-k}) e^{-\left(\frac{k-b_{ij}^{+}}{c_{ij}^{+}}\right)^2}, \ \forall k > 0, \varepsilon_{t-k} > 0 \\ a_{ij}^{+}(\varepsilon_{t-k}) = \alpha_{ij}^{+} + \beta_{ij}^{+} |\varepsilon_{t-k}| \end{cases} \tag{11}$$

with α_{ij}^{+}, β_{ij}^{+}, b_{ij}^{+} and c_{ij}^{+} parameters to be estimated. An identical functional form holds for $\psi_{k,ij}^{-}$.

In this model, the maximum effect of a shock on a given variable can depend on the size of the shock, and that relation may be different for positive and negative shocks. In other words, the marginal effect of a monetary shock is allowed to depend on both the value and the sign of the shock. In (11), we allow the scaling factor of the impulse response (a_{ij}) to be a linear function of the value of the shock.

4.6 Asymmetry, size-dependence and state dependence

Finally, we can also consider a more general framework in which the effect of a shock also depends on the state of the business cycle.

Denoting z_t a business cycle indicator variable that is a function of past observables $\mathbf{Z}_t$ (e.g., lagged output growth or output growth over the past X months), and using a one-Gaussian approximation, the impulse response function following a positive innovation can be parametrized as

$$\begin{cases} \psi_{ij}^{+}(k) = a_{ij}^{+}(\varepsilon_{t-k}) e^{-\left(\frac{k-b_{ij}^{+}}{c_{ij}^{+}}\right)^2}, \ \forall k > 0, \varepsilon_{t-k} > 0 \\ a_{ij}^{+}(\varepsilon_{t-k}) = \alpha_{ij}^{+} + \beta_{ij}^{+} |\varepsilon_{t-k}| + \gamma_{ij}^{+} z_{t-k} \end{cases} \tag{12}$$

with α_{ij}^{+}, β_{ij}^{+}, γ_{ij}^{+}, b_{ij}^{+} and c_{ij}^{+} parameters to be estimated. An identical functional form holds for $\psi_{k,ij}^{-}$.

In this model, the amplitude of the impulse response depends on the sign and size of the shock, as well as the state of the business cycle at the time of the shock. In (12), we allow the amplitude of the impulse response (a_{ij}) to be a linear function of the indicator variable z_t. Such a specification allows us to test whether, for instance, an expansionary policy has a positive effect on output in recessions than in expansions.

5 Estimation

Once parametrized, model (2) can be estimated by maximum likelihood. In this section, we first describe how we construct the likelihood function by exploiting the prediction-error

decomposition, and then we discuss the estimation routine.

5.1 Constructing the conditional likelihood function

We want to build the likelihood function for a sample of size T and parameter vector θ: $p(\mathbf{Z}_1, ..., \mathbf{Z}_T|\theta)$. To do this, we decompose the likelihood using:[21]

$$p(\mathbf{Z}_1, ..., \mathbf{Z}_T|\theta) = p(\mathbf{Z}_T|\mathbf{Z}_1, ..., \mathbf{Z}_{T-1}, \theta) \dots p(\mathbf{Z}_2|\mathbf{Z}_1, \theta)p(\mathbf{Z}_1|\theta) \tag{13}$$

To calculate the one-step-ahead conditional likelihood function needed for the prediction error decomposition, we proceed as follows. First, we assume that all innovations $\{\varepsilon_t\}$ are Gaussian with mean zero and variance one.[22] We then use the MA representation (2) and the recursive identification scheme to uniquely identify the current period value of the innovations as a function of current period data and past innovations.

Specifically, truncating the MA(∞) representation at some level K, we set the first K values for ε to the zero vector.[23] Conditional on knowing the first $[t-K, t]$ innovations, the $t+1$ innovation is obtained from:

$$\mathbf{\Psi}_0^{\pm}(\varepsilon_{t+1})\,\varepsilon_{t+1} = \mathbf{Z}_t - \sum_{k=0}^{K} \mathbf{\Psi}_k^{+} \mathbf{1}_{\varepsilon \geq \mathbf{0}} \varepsilon_{t-k} + \sum_{k=0}^{K} \mathbf{\Psi}_k^{-} \mathbf{1}_{\varepsilon < \mathbf{0}} \varepsilon_{t-k}$$

with $\mathbf{\Psi}_0^{\pm}(\varepsilon_{t+1})$ a function of ε_{t+1} since the elements of $\mathbf{\Psi}_0^{\pm}$, the contemporaneous lag parameter matrix, depends on the sign of the elements of the innovation vector ε_{t+1}.

To calculate ε_{t+1} as a function of the parameters, one needs to uniquely pin down the value and sign of the components of ε_{t+1} as a function of the parameters. Fortunately, this is straightforward with a recursive identifying restriction scheme, since $\Psi_0^{\pm}(\varepsilon_{t+1})$ is lower triangular.

Denoting $\psi_{0,ij}^{+}$ the element of Ψ_0^{+} and using the bivariate case as an illustration, we have

$$\Psi_0^{\pm}(\varepsilon_{t+1}) = \begin{pmatrix} \psi_{0,11}^{\pm}(\varepsilon_{1,t+1}) & 0 \\ \psi_{0,21}^{\pm}(\varepsilon_{1,t+1}) & \psi_{0,22}^{\pm}(\varepsilon_{2,t+1}) \end{pmatrix} \tag{14}$$

with $\psi_{0,11}^{\pm}(\varepsilon_{1,t+1})$ indicating that the value of $\psi_{0,11}^{\pm}$ depends on $\varepsilon_{1,t+1}$, $\psi_{0,22}^{\pm}(\varepsilon_{2,t+1})$ depends on

[21] To derive the conditional densities in decomposition 13, our parameter vector θ thus implicitly also includes initial values for the lagged innovations. We will keep those fixed throughout the estimation and discuss alternative initializations below.

[22] The estimation could easily be generalized to allow for non-normal innovations such as t-distributed errors. Note that even with Gaussian structural innovations, our non-linear model implies that the reduced-form residuals need not be Gaussian.

[23] Alternatively, one can use the first K values of the shocks recovered using an SVAR.

$\varepsilon_{2,t+1}$, etc... A similar form for would hold in a higher-dimensional case.

The fact that $\Psi_0^{\pm}(\varepsilon_{t+1})$ is lower triangular with positive entries on the diagonal allows us to uniquely identify the value (and sign) of $\varepsilon_{1,t+1}$, the first component of ε_{t+1}, from the first equation with

$$\varepsilon_{1,t+1} = \frac{\mathbf{Z}_{1,t}}{\psi_{0,11}^{\pm}(\varepsilon_{1,t+1})}$$

and then uniquely identify the sign and value of each element of ε_t by iterating on (2). For instance, for $\varepsilon_{2,t+1}$, we have

$$\varepsilon_{2,t+1} = \frac{\mathbf{Z}_{2,t} - \psi_{0,21}^{\pm}(\varepsilon_{1,t+1})\varepsilon_{1,t+1}}{\psi_{0,22}^{\pm}(\varepsilon_{2,t+1})}. \tag{15}$$

Looking at (15), once the sign and value of $\varepsilon_{1,t+1}$ is known, the first column of $\Psi_0^{\pm}$ is pinned down (and thus $\psi_{0,21}^{\pm}(\varepsilon_{1,t+1})$ is known). Since the only unknown left in (15) is $\psi_{0,22}^{\pm}$, which is positive, we know the sign of $\varepsilon_{2,t+1}$ from the sign of $\mathbf{Z}_{2,t} - \psi_{0,21}^{\pm}(\varepsilon_{1,t+1})\varepsilon_{1,t+1}$. We then get the value of $\varepsilon_{2,t+1}$ from (15).

In a higher-dimensional case, we can proceed identically by iterating using (2) to uniquely identify the sign and value of each element of ε_t.

With the series of $\{\varepsilon_t\}$ in hand, we can then construct the conditional log likelihood function from equation (2).

When the model allows for state dependence, as in specification (12), the likelihood also depends on the value of the indicator function z_t. Technically, constructing the likelihood of this specification is a straightforward extension of our previous cases, because z_t is a function of lagged values of Y_t. To see that, note that we use the prediction-error decomposition to construct the likelihood function. Thus, we build a sequence of densities for z_t that condition on past values of z_t, which includes past values of Y_t. Thus, conditional on past values of Y_t, z_t is known.[24]

5.2 Estimation routine

We now discuss the estimation routine. To estimate our model, we use the Metropolis-Hastings algorithm (see Robert & Casella 2004, Haario et al., 2001). To be specific, we use a Metropolis-within-Gibbs algorithm with the blocks given by the different groups of parameters in our model

[24] If we wanted to use an indicator function that was not part of the vector of endogenous variables Y_t, this would also be possible by using a quasi-likelihood approach. That is, we would build a likelihood function that not only conditions on the parameters, but also the sequence of indicators z_t. This would in general not be efficient because the joint density of z_t and Y_t could carry more information about the parameters in our model than the conditional density we advocate using. As long as z_t is highly correlated with elements of (functions of) Y_t, this loss in efficacy will likely be small.

(one block being composed of all a parameters, another composed of all b parameters and so on). As described above, we can build the likelihood once we assume a distribution for the innovations in our model. We use improper uniform priors so that our results are interpretable as maximum-likelihood estimates.

Importantly, it is straightforward to incorporate prior information in the Metropolis-Hastings algorithm, and our approach makes prior elicitation easy, because our coefficients are directly interpretable as impulse response coefficients.

Since we have a relatively large parameter space, it is important to initialize the Metropolis-Hastings algorithm in an area of the parameter space that has substantial posterior probability. To do so, we follow a two-step procedure: first, we estimate a standard VAR using OLS on our data set, calculate the MA representation, and we use the impulse response functions implied by the VAR as our starting point. Specifically, we first fit our parametrized impulse response functions to the VAR-based impulse response functions.[25] We then use these coefficients as a starting point for our maximization routines[26] that then give us a starting value for the Metropolis-Hastings algorithm.

Finally, in order to reduce the parameter space, in our benchmark estimations, we only allow for asymmetry or non-linearity in the impulse response functions to monetary shocks. The impulse response functions to the other shocks are set at their values estimated from the corresponding SVAR. However, as a robustness check, we can verify that our findings are not driven by this restriction, by estimating a model with non-linearity in response to all shocks.

6 Measuring the asymmetric effects of monetary shocks

In this section, we present our empirical results on the asymmetric effect of monetary shocks. After validating our parametric estimation procedure in the symmetric case using the standard SVAR approach as a benchmark, we present the asymmetric impulse response functions.

Our benchmark analysis is based on monthly data covering the period 1959:M1 to 2006:M12. Following the literature (e.g., Bernanke and Blinder, 1992), we take the federal funds rate as our measure of policy, and use innovations in the federal funds rate as a measure of monetary policy shocks. This is the reason why we exclude the 2008-2008 recession from our analysis, since the fed funds rate stood at the zero-lower bound for the whole period and no longer captured variations in the stance of monetary policy. We consider a standard trivariate monthly model with the log of industrial production, the PCE inflation rate and the federal funds

[25] Specifically, we set the parameters of our model (for instance, a, b and c in the benchmark model with one Gaussian function) to minimize the discrepancy (sum of squared residuals) between the two sets of impulse responses.

[26] We use both a simplex algorithm and Chris Sims' csminwel routine.

rate. All series were detrended with a quartic trend.[27] When constructing the likelihood, we truncate the MA representation at $K = 80$ months.

6.1 The linear case: VAR versus parametric estimation

First, as a validation of our proposed approach to directly estimate impulse-responses parametrically, we evaluate how our parametric approach, either using a one-Gaussian approximation or a two-Gaussian approximation performs in fitting the impulse response functions implied by a standard VAR with a recursive identifying restriction.

As previously discussed, a one-Gaussian approximation does a very good job at capturing the impulse response functions to a monetary shock implied by a VAR (Figure 1), only missing the initially slow response of output and the price puzzle.[28] Using two Gaussians instead of one corrects this problem. As shown in Figure 3, a two-Gaussian approximation does an excellent job at capturing the impulse response functions to a monetary shock implied by a VAR, capturing both the initially slow response of output as well as the price puzzle.[29]

Table 1 compares the log-likelihood of the VAR model with that from our parametric model, first using a one-Gaussian approximation and second using a two-Gaussian approximation. We can see that our parametric approach fits the data almost as well as the VAR.[30]

6.2 Evidence for asymmetry

We now estimate our model allowing for asymmetry depending on the sign of the monetary shock. In our benchmark specification, we parameterize the impulse response functions with one-Gaussian approximation.

Allowing for asymmetry provides a dramatic improvement in the model fit of the data. As shown in Table 1, the log-likelihood is about 150 log-points higher when allowing for asymmetry. Looking at the impulse responses to a monetary shock, Figure 4 plots the responses (in percentage points) of output (industrial production), the price level and the federal funds rate.

The evidence for asymmetry is striking: following a positive shock to the fed funds rate (a contractionary policy), output growth decreases markedly. In contrast, following a negative shock to the fed funds rate (an expansionary policy), the response of unemployment is small

[27] Alternative filters such as a conventional HP filter give very similar results.

[28] The specification of the Gaussian approximation (the parameters a, b, and c) is found by minimizing the distance between the VAR-based impulse response and the parametric impulse response.

[29] The specification of the two-Gaussian approximation (the parameters a_l, b_l, c_l, $l = \{1, 2\}$) is found by minimizing the distance between the VAR-based impulse response and the parametric impulse response.

[30] Note that the VAR has more free parameters than the parametrized model. A VAR specified with 12 lags (as in this specification) has $3 * 12 + 3 = 39$ free parameters, while the one-Gaussian model has only $3^3 + 3 = 30$ free parameters.

and non-significantly different from zero. This evidence suggests that the contractionary effect of monetary policy may have been substantially under-estimated, as the adverse effect of a contractionary monetary shock on output is as much as two- to three-times larger than suggested by the VAR estimate.

Importantly, note that the responses of the federal funds rate are very similar after positive or negative shocks, with similar persistence levels, indicating that the asymmetric responses of output are not driven by differences in the responses of the fed funds rate. The response of prices is also somewhat asymmetric: the price level declines more following a contractionary shock than it increases following an expansionary shock.

To assess whether the impulse response functions to positive and negative monetary shocks are statistically significantly different, we consider the difference, in absolute value, between the peak of the impulse response function to a positive shock and the peak of the impulse response function to a negative shock of the same size, i.e.,

$$D = \sup_k \left|\psi_k^+\right| - \sup_k \left|\psi_k^-\right|$$

where ψ_k^+ denote the period k response to a positive shock and ψ_k^- the period k response to a negative shock. With the one-Gaussian approximation, the D statistics is simply

$$D = a^+ - a^-.$$

Table 2 shows that, for output, the D-statistic is large and significantly different from zero, as the difference between the peak effect of a positive and the peak effect of a negative shock is large; in the order of magnitude of the peak response estimated from a VAR.

6.3 Robustness to model specification

In this subsection, we present results on the robustness of our baseline specification with asymmetry.

6.3.1 Two-Gaussian parametrization

To allow for a richer set of impulse-responses, in particular to allow for an overshooting (or oscillating) pattern in the impulse-response, we consider a two-Gaussian approximation of the impulse response functions, following (6).

Figure 5 plots the estimated impulse-responses to a monetary shock and points to the same conclusion: a contractionary monetary shock has a strong effect, while expansionary shocks

have no significant effect. The response of the price level is in line with that of the one-Gaussian specification: inflation's response is stronger following a contractionary shock.

6.3.2 Post-1982 period

The 1979-1982 period was a time of unusual monetary policy, as the Fed officially stopped targeting the Fed funds rate. Although this period may also be particularly instructive for identifying the effects of monetary shocks, displaying the largest monetary shocks in the sample and being a clear example of a monetary policy driven recession (Coibion, 2012), we assess the robustness of our results to this period by estimating our model using data covering only 1983-2006.

Evidence for asymmetry is still strongly present: Table 2 shows that the D-statistic, although smaller than with the full sample and less precisely estimated (given that the sample size is less than one half of the original one), is still large (in the order of magnitude of the peak response estimated from a VAR) and significantly different from 0.

6.3.3 Alternative specification

We verified the robustness of our results to using the growth rate of industrial production, instead of the (log) level, with inflation and the fed funds rate as used as previously. In this specification, the data were not previously filtered. Figure 7 shows again strong evidence that monetary policy has asymmetric effects on output.

Finally, using the unemployment rate as our measure of real activity and the GDP deflator as our price measure does not change our conclusions.[31] The impulse response function (not shown) are qualitatively similar, and Table 2 shows that the D statistic is large (in the order of magnitude of the peak response estimated from a VAR) and significantly different from zero both when using the whole 1959-2006 sample period or the 1983-2006 period.

6.3.4 Allowing for asymmetry in response to all shocks

Finally, although the baseline specifications only allowed for asymmetry in response to monetary shocks, and not in response to the other (unidentified) shocks, it is also possible to allow for asymmetry with respect to the other disturbances. Although it considerably increases the parameter space, we find evidence for asymmetry in response to monetary shocks but no evidence for asymmetry in response to the non-monetary shocks, in line with our benchmark approach. As shown in Table 2, the D-statistics is highly significant, and in line with the findings from previous specifications.

[31] The GDP deflator was linearly interpolated at a monthly frequency from quarterly data.

6.4 Robustness to shock identification: using the Romer and Romer monetary shocks

Our evidence for asymmetry has so far been based on using a recursive identifying restriction to recover the monetary shock. In this section, we instead use monetary shocks identified through a different approach: the narrative approach from Romer and Romer (2004). Proceeding as in Romer and Romer (2004), we use the Romer and Romer monetary shocks as a measure of the stance of monetary policy in our model (2).[32]

As pointed out by Coibion (2012), the advantage of this procedure is that one should be able to more precisely identify the effects of monetary shocks than in the previous specification with the fed funds rate, since the Romer and Romer measure controls for much of the endogenous fluctuations in the interest rate as well as the Fed's information set.

The Romer and Romer series were recently extended by Coibion et al. (2012), so that our sample covers 1966-2006. Figure 6 plots the results and shows similar results as previously with strong evidence that monetary policy has asymmetric effects on output. And again, the results suggest that the contractionary effect of monetary policy on output is substantially larger than suggested by VAR estimates.

7 Size dependence and the effects of monetary policy

In this section, we enrich our empirical model by allowing for size-dependence in the effect of monetary shocks, that is by allowing the effect of monetary policy to depend on the magnitude of the intervention. We do so in a simple way; by allowing the peak effect of monetary policy (the a coefficient in the in the one-Gaussian case) to depend on the size of the shock. Given the strong evidence for asymmetry found in the previous section, we still allow for asymmetric effects by monetary policy, so that we estimate specification (11) described in Section 4.

Figure 8 presents the estimated impulse response functions as we vary the size of the monetary shock from -1.5 standard deviation to +1.5 standard deviation (roughly .70 percentage points). While there appears to be little non-linear effect of shock size on the impulse responses of inflation and the fed funds rate, the response of output seems to depend non-linearly on the size of the shock.

To visualize this result more easily, Figure 9 plots the peak effects of a monetary shock ε^p on output, $-\max_k \mathbf{\Psi}_k(\varepsilon^p)\varepsilon^p-$, as a function of the value of the monetary shock ε^p. As discussed previously, with a one-Gaussian parametrization, the peak effect is given straightforwardly by $a_{ij}^{\pm}(\varepsilon^p)\varepsilon^p$. Figure 9 plots the peak effects on output growth for shocks varying from -2 to $+2$

[32] Alternatively, we could have used the cumulative sum of Romer and Romer monetary shocks as the measure of monetary policy. Results are very similar.

standard deviations, which correspond to shocks ranging -1.1 to +.85 percentage points of fed funds rate.[33] To ease comparison with estimates from the corresponding linear VAR, Figure 9 also shows (thick dashed black line) the peak effect of a monetary shock as a function of the size of the shock. Obviously, that effect is linear.

It is clear from Figure 9 that monetary shocks have asymmetric effects on output, as previously discussed. Moreover, the results show that the VAR estimates lie in between the estimates for expansionary and contractionary shocks. According to our estimates, the effects of contractionary shocks seem substantially understated by the VAR while the effects of expansionary shocks seem substantially overstated.

Turning to size dependence, the non-linearities work in opposite directions for positive and negative shocks. While small expansionary shocks are clearly neutral, large expansionary shocks may have some positive effect on output, although the error bands are large. In fact, an expansionary shock can have some positive effect on output (although still smaller than suggested by a VAR estimate), but the policy intervention has to be large (a cut in the fed funds rate of at least .50 percentage points).

Interestingly, the opposite is true for contractionary shocks. Small contractionary shocks have the strongest marginal effect on output, and as one increases the size of the contractionary shock, the marginal effect of the shock slightly declines. In fact, for shocks above +.85 percentage points, the marginal effect of the shock is equal to the VAR estimate. Importantly, this finding indicates that the strong contractionary effect of monetary policy that we uncovered is not driven by large outlier shocks.

8 State dependence and the effects on monetary policy

We now further enrich our model by allowing the effect of monetary policy to depend on the state of the business cycle. Intuitively, we would like to test whether monetary policy is more powerful at stimulating the economy in a period of low growth than in a period of high growth, and whether the response of inflation depends on the amount of slack in the economy.

We thus estimate specification (12), where the peak effect of monetary policy is allowed to depend on (i) the sign of the shock, (ii) the magnitude of the shock, and (iii) the state of the business cycle at the time of the shock. As the cyclical indicator z_t, we use the average monthly growth rate in industrial production growth over the past 12 months.[34]

Taken together, our results show that, during recessionary periods, expansionary monetary

[33] Variation across shock sizes is asymmetric, because the variance of expansionary shocks was estimated to be larger than the variance of contractionary shocks.

[34] Using alternative definitions of the cyclical indicator, such as the IP growth rate over the past month or past 6 months give very similar results.

policy does have some stimulating effect on output with no inflationary consequences, but that, during expansionary periods, expansionary monetary policy only generates inflation. The neutral effect of expansionary policy uncovered previously (Figure 4 or 9) thus seems to be the result of a "composition effect" coming from the fact that the economy alternates between expansions and recessions. Monetary policy does have some simulating effects on output, but *only* in recessions. As a result, on average, the expansionary effect appears uncertain and not significantly different from zero.

Before discussing our parameter estimates in more details, Table 1 shows that the log-likelihood of the state-dependent model is substantially higher than a simpler model with only asymmetry and size-dependence, and a likelihood-ratio test easily reject the null in favor of the model with state-dependence.

As a preliminary step, and to put our results into perspective, Figure 10 plots the annualized monthly growth rate of IP (i.e., the indicator variable z_t) over the past 50 years, along with the identified monetary shocks, and Figure 11 plots the distribution of estimated monetary shocks over the state of the business cycle. We can see that monetary shocks are roughly equally distributed between expansions and recessions, and that shocks are observed over a wide range of values of the indicator function, increasing our confidence in the validity of our estimates.

To visualize our estimates, Figure 12 plots how the peak effects of monetary policy on output and inflation vary with the state of the cycle, both for contractionary and for expansionary shocks.[35] In all cases, we explore the responses of output and inflation to a one standard-deviation monetary policy shock, which corresponds to a 45 basis points change in the fed funds rate.

Starting with output, depicted in the left-hand column of Figure 12, we can first notice that the effect of a contractionary policy (top left quadrant) is relatively independent of the state of the business cycle, always having a large contractionary effect on output and substantially larger than estimated with a VAR. A very different picture holds for expansionary shocks (bottom left quadrant). This time, the effect of monetary policy depends crucially on the state of the business cycle: the lower the level of output (IP) growth, the larger the effect of an expansionary policy. In fact, it is only in large slumps that monetary policy can have any expansionary effect on output. In fact, the effect becomes significantly positive when (annualized) monthly IP growth falls to about -4.5%. Figure 10 shows that growth rates in IP of -4.5% are not uncommon, having been observed in all past NBER recessions.[36] Thus, expansionary policy does seem to have some effect in periods of great slack. However, the asymmetry

[35] Specifically, the slopes of the solid lines depicted in Figure 12 are the estimated γ^+ and γ^- coefficients for output and inflation.

[36] While the unconditional average IP growth rate over 1959-2006 is about 3.2%, the average IP growth during NBER recessions is roughly -3%.

between expansionary and contractionary interventions remains, and an expansionary policy is always considerably less potent than its contractionary counterpart. For instance, even in a deep recession (IP grwoth rate of 4.5%), a 45 basis point rate cut has a substantially smaller effect than the corresponding rate increase (about +0.3 percentage points increase in output versus a -0.8 ppt decline).

We now turn to the response of inflation, depicted in the right-hand column of Figure 12. Although the estimates are more uncertain, the estimates suggest that, regardless of the sign of the shock, the stronger the state of the economy, the more responsive is inflation to monetary shocks.[37] In particular, while an expansionary policy has no inflationary effect in a recession, it does raise prices in an expansion, although again the point estimates are more uncertain. A similar conclusion holds for contractionary shocks; a contractionary shock has a stronger deflationary effect in a boom, with the point estimate being significantly different from zero when the economy is growing.

To look at our estimates from a different angle, we now keep the state of the business cycle constant but vary the size of the monetary shock, as in the previous section, for two opposite phases of the business cycle: an expansionary stage and a recessionary stage. We call an expansionary phase a period where IP growth is one standard-deviation above its mean (annualized growth rate of about +8%), and we call a recessionary phase a period where IP growth is one standard-deviation below its mean (growth rate of about -2%).

Similarly to Figure 9, Figure 13 and Figure 14 plot the peak effect of monetary policy on output and inflation as a function of the size of the shock, varying the size of the monetary intervention from -2 to +2 standard deviations, which correspond to shocks ranging -1.1 to +.85 percentage points of fed funds rate.[38] Comparing the top and bottom panels of Figure 13 clearly shows how expansionary monetary policy only has expansionary effects on output in recessions. And supporting our previous conclusion on size-dependence, the larger the expansionary shock, the stronger the expansionary effect of policy on output. In expansions however, expansionary policy is close to neutral, and only lead to inflation (Figure 14). In fact, the inflationary effect of an expansionary policy during a boom increases more than proportionally with the size of the shock. For an 80 basis points reduction in the fed funds rate during an expansion, the inflationary effect is already twice as large as that implied by the VAR estimate.

In contrast, the effect of contractionary shocks are relatively insensitive to the state of the

[37] The γ^+ and γ^- coefficients for inflation are not statistically different from each other, meaning that the effect of the output gap on the effect of monetary policy on inflation is independent on the sign of the monetary intervention. This is not the case for the γ^+ and γ^- coefficients for output.

[38] Again, the variation in shock size is asymmetric, because the variance of expansionary shocks was estimated to be larger than the variance of contractionary shocks.

business cycle. The effect on output are large, and much larger than estimated with a linear VAR, while the effects on inflation are not significantly different from the VAR estimates.

Finally, as a robustness check, we repeat our exercise using the Romer and Romer monetary shocks to capture the stance of monetary policy, instead of the fed funds rate, as in sub-section 6.4. Similarly to Figure 12, Figure 15 plots how the peak effects of monetary policy on output and inflation vary with the state of the cycle, both for contractionary and for expansionary shocks.

Our main result that an expansionary policy can only have an expansionary effect during recessions is still clearly present, with a similar point estimate for the γ^- coefficient for output (the slope of the solid line in the bottom left quadrant). The effect of the state of the cycle on the response of inflation is also qualitatively similar, although the point estimates are smaller, meaning that the response of inflation appears less sensitive to the state of the cycle than previously estimated. Finally, the contractionary effect of policy on output (upper left quadrant) now seems to depend on the state of the cycle with a contractionary policy having a stronger adverse effect on output in a recession, although looking at the error bands, one cannot reject the null of no state dependence at the 90% confidence level.

9 Conclusion

This paper provides new evidence for the existence of asymmetry, size-dependence and state-dependence in the transmission mechanism of monetary policy. To do so, we develop a non-linear method to identify the effects of policy shocks, which consists in parameterizing the impulse response functions using Gaussian basis functions.

We find that the effect of a monetary shock depends strongly on (i) the sign on the shock, (ii) the size of the shock, and (iii) the state of the business cycle at the time of the shock. A contractionary policy has a strong adverse effect on output, much stronger than linear estimates suggest, but an expansionary policy has, on average, no significant effect on output. An expansionary policy can have some expansionary effect on output, but only if the intervention is large and during a recession. However, an expansionary policy is always considerably less potent than its contractionary counterpart.

Although we focused only on monetary policy in this paper, our novel parametric approach could be applied to other recursive identification schemes and could thus be used to estimate the non-linear effects of other important shocks where the existence of asymmetry, size-dependence or state-dependence remains an important and unresolved question; for instance, fiscal policy shocks (Romer and Romer, 2010, Auerbach and Gorodnichenko, 2012, Ramey and Zubairy, 2014), oil price shocks (Hamilton, 2003) or credit spread shocks (Gilchrist and Zakrajsek,

2012).

References

[1] Almon, S. "The distributed lag between capital appropriations and expenditures," Econometrica, 33, January, 178-196, 1965

[2] Angrist, J., Jorda O and G. Kuersteiner. "Semiparametric estimates of monetary policy effects: string theory revisited," NBER Working Paper, 2013

[3] Alspachs D. and H. Sorenson H. "Recursive Bayesian Estimation Using Gaussian Sums," *Automatica*, Vol 7, pp 465-479, 1971.

[4] Alspachs D. and H. Sorenson H. "Nonlinear Bayesian Estimation Using Gaussian Sum Approximations," *IEEE Transactions on Automatic Control*, Vol 17-4, August 1972.

[5] Auerbach A and Y Gorodnichenko. "Measuring the Output Responses to Fiscal Policy," American Economic Journal: Economic Policy, vol. 4(2), pages 1-27, 2012

[6] Ball L. and G. Mankiw. "Asymmetric Price Adjustment and Economic Fluctuations," Economic Journal, 104(423), pages 247-61, 1994

[7] Ball, L. and D. Romer, "Real Rigidities and the Non-neutrality of Money," Review of Economic Studies, Wiley Blackwell, vol. 57(2), pages 183-203, April 1990.

[8] Barro, Robert J., "Unanticipated Money Growth and Unemployment in the United States," American Economic Review, LXVII, 101-15, 1977.

[9] Beaudry, Paul, and Gary Koop. "Do Recessions Permanently Change Output?" Journal of Monetary Economics 31 (1993), 149 - 63.

[10] Buhmann, Martin D, Radial Basis Functions: Theory and Implementations, Cambridge University Press, 2003.

[11] DeLong B and L. Summers, "How Does Macroeconomic Policy Affect Output?," Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, vol. 19(2), pages 433-494, 1988.

[12] Cover, J. "Asymmetric Effects of Positive and Negative Money-Supply Shocks," The Quarterly Journal of Economics, Vol. 107, No. 4, pp. 1261-1282, 1992.

[13] Coibion, O. "Are the Effects of Monetary Policy Shocks Big or Small?," American Economic Journal: Macroeconomics, vol. 4(2), pages 1-32, April, 2012.

[14] Coibion, O., Y. Gorodnichenko, L. Kueng, and J. Silvia, "Innocent Bystanders? Monetary Policy and Inequality in the US," NBER Working Papers 18170, National Bureau of Economic Research, Inc, 2012.

[15] Christiano, L., M. Eichenbaum, and C. Evans. "Monetary policy shocks: What have we learned and to what end?," Handbook of Macroeconomics, volume 1, chapter 2, pages 65-148 Elsevier, 1999.

[16] Gilchrist, S and E Zakrajsek, "Credit Spreads and Business Cycle Fluctuations," American Economic Review, American Economic Association, vol. 102(4), pages 1692-1720, June 2012.

[17] Koop G, M. Pesaran and S. Potter "Impulse response analysis in nonliner multivariate models," Journal of Econometrics, 74 119-147, 1996.

[18] Koop G and S. Potter "Dynamic asymmetries in US unemployment," Journal of Business & Economic Statistics Volume 17, Issue 3, 1999

[19] Koreyaar J. *Mathematical Methods*, Vol. 1, pp 330-333. Academic Press, New York, 1968.

[20] McLachlan, G. and D. Peel. *Finite Mixture Models*. Wiley Series in Probability and Statistics, 2000.

[21] Morgan, D. "Asymmetric Effects of Monetary Policy." Federal Reserve Bank of Kansas City Economic Review 78, 21-33, 1993

[22] Olivei G. and S. Tenreyro, "The Timing of Monetary Policy Shocks," American Economic Review, vol. 97(3), pages 636-663, 2007.

[23] Potter S "Anonlinear approach to US GNP," Journal of Applied Econometrics," Vol 10 109-125, 1995

[24] Robert, Christian P. and George Casella "Monte Carlo Satistical Methods" Springer, 2004.

[25] Ravn M and M Sola. "A Reconsideration of the Empirical Evidence on the Asymmetric Effects of Money-supply shocks: Positive vs. Negative or Big vs. Small," Archive Discussion Papers 9606, Birkbeck, 1996.

[26] Ravn M. and M. Sola. "Asymmetric effects of monetary policy in the United States," Review, Federal Reserve Bank of St. Louis, issue Sep, pages 41-60, 2004.

[27] Ramey V. and S. Zubairy. "Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data ," Working Paper, 2014.

[28] Romer, C., and D. Romer. “A New Measure of Monetary Shocks: Derivation and Implications,” American Economic Review 94 (4): 1055–84, 2004

[29] Romer, C., and D. Romer. “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks,” American Economic Review, June 2010

[30] Tenreyro, S., and G. Thwaites (2013): “Pushing on a string: US monetary policy is less powerful in recessions,” Working Paper.

[31] Thoma, M. ”Subsample Instability and Asymmetries in Money-Income Causality.” Journal of Econometrics 64 (1994), 279-306.

[32] Uhlig, H. ”What are the effects of monetary policy on output? Results from an agnostic identification procedure,” Journal of Monetary Economics, Elsevier, vol. 52(2), pages 381-419, March 2005.

[33] Walsh C. Monetary Theory and Policy, 3nd. ed., The MIT Press, 2010.

[34] Weise, C. ”The Asymmetric Effects of Monetary Policy: A Nonlinear Vector Autoregression Approach,” Journal of Money, Credit and Banking, Blackwell Publishing, vol. 31(1), pages 85-108, February 1999.

Appendix: Proof of Theorem 1

First, we split the continuous function $f(x)$ over subintervals for which the sign of $f(x)$ is constant. Specifically, since $f(x)$ is continuous, there exists an integer K and $\{x_1, x_2...x_K\}$ in $[0, \infty[$ such that $\forall x \in [x_k, x_{k+1}]$ $k \in \{1, K\}$, $f(x) \leq 0$ or $f(x) \geq 0$.

From now on, we will consider the problem of approximating the function $f(x) \geq 0$ over the interval $[x_k, x_{k+1}]$. This is without loss of generality, because if $f(x) \leq 0$, we can instead consider the function $-f(x)$.

Following Alspach and Sorenson (1971, 1972) in the context of approximating distributions, the problem of approximating $f(x) \geq 0$ over $[x_k, x_{k+1}]$ can be considered within the context of delta families of positive types. Basically, delta families are families of functions which converge to a delta function as a parameter characterizing the family converges to a limit value.

Let $\delta_\lambda(x)$ be a family of functions on the interval $]-\infty, +\infty[$ which are integrable over every interval. This is called a delta family of positive type is the following conditions are satisfied

1. $\int_{-s}^{s} \delta_\lambda(x)dx \longrightarrow 1$ as λ tends to some limit value λ_0 for some s.
2. For every constant $\gamma > 0$, δ_λ tends to zero uniformly for $\gamma \leq |x| \leq \infty$ as $\lambda \to \lambda_0$
3. $\delta_\lambda(x) \geq 0$ for all x and λ.

Using the delta families, the following result can be used for the approximation of a function $f(x)$:

Theorem: The sequence $f_\lambda(x)$ which is formed by the convolution of δ_λ and f

$$f_\lambda(x) = \int_{-\infty}^{+\infty} \delta_\lambda(x-u)f(u)du \tag{16}$$

converges uniformly to $f(x)$ as $\lambda \to \lambda_0$ on every interior subinterval of $]-\infty, +\infty[$.

Proof: see Korevaar 1968.

This result can be applied to Gaussian functions by noting that Gaussian functions form a family of delta family of positive types. Defining

$$\delta_\lambda(x) \equiv G_\lambda(x) = \frac{1}{\sqrt{2\pi\lambda^2}} e^{-\frac{x^2}{\lambda^2}}, \tag{17}$$

it is easy to show that the Gaussian function $G_\lambda(x)$ forms a delta family of positive type as $\lambda \to 0$ (i.e., $\lambda_0 = 0$). That is, the Gaussian function tends to the delta function as the variance tends to zero.

Using (17) in (16), the function $f_\lambda(x)$ converges uniformly to $f(x)$ as $\lambda \to 0$ where $f_\lambda(x)$ is

given by

$$f_\lambda(x) = \int_{-\infty}^{+\infty} G_\lambda(x-u) f(u) du, \tag{18}$$

which forms the basis of the Gaussian sum approximation.

Since the term $G_\lambda(x-u)f(u)$ is integrable over $]-\infty, +\infty[$ and continuous, (18) can itself be approximated by a Riemann sum. In particular, an approximation of $f_\lambda(x)$ over $[x_k,\, x_{k+1}]$ of $f(x)$ is given by

$$f_{\lambda,N}(x) = \sum_{n=1}^{N} f(x_n) G_\lambda(x-x_n)\,(\xi_n - \xi_{n-1})$$

where the interval $[x_k,\, x_{k+1}]$ is divided into N subintervals by selecting points ξ_n such

$$x_k = \xi_0 < \xi_1 < ...\xi_N = x_{k+1}$$

and where in each subinterval $[\xi_k,\, \xi_{k+1}]$, the point x_n is chosen such that

$$f(x_n)\,(\xi_n - \xi_{n-1}) = \int_{\xi_{n-1}}^{\xi_n} f(x) dx$$

which is possible by the mean value theorem.

Defining $\lambda = \frac{c_0}{N}$ with c_0 a constant, an approximation of $f(x)$ over $[x_k,\, x_{k+1}]$ can be constructed from $f_N(x)$; a combination of Gaussian functions with

$$f_N(x) \equiv \sum_{n=1}^{N} \alpha_n G_\lambda(x-x_n) \tag{19}$$

with $\alpha_n = f(x_n)\,(\xi_n - \xi_{n-1})$.

The final step of the proof is to approximate $f(x)$ over any arbitrary interior interval $[a,b]$. By dividing the interval $[a,b]$ into subintervals $I_k = [x_k, x_{k+1}]$ over which $f(x)$ has constant sign, we can apply results (19) on each subinterval such that the function $f(x)$ can be approximated by

$$f_N(x) = \sum_{n=1}^{N} \sum_{I_k} \alpha_n^{I_k} G_\lambda(x - x_n^{I_k}) \tag{20}$$

with $\alpha_n^{I_k}$ and $x_n^{I_k}$ defined as previously over $I_k = [x_k, x_{k+1}]$.

Defining $a_n = \alpha_n^{I_k}$ and $b_n = x_n^{I_k}$, we get $f_N(x) \underset{N\to\infty}{\longrightarrow} f(x)$ which completes the proof.

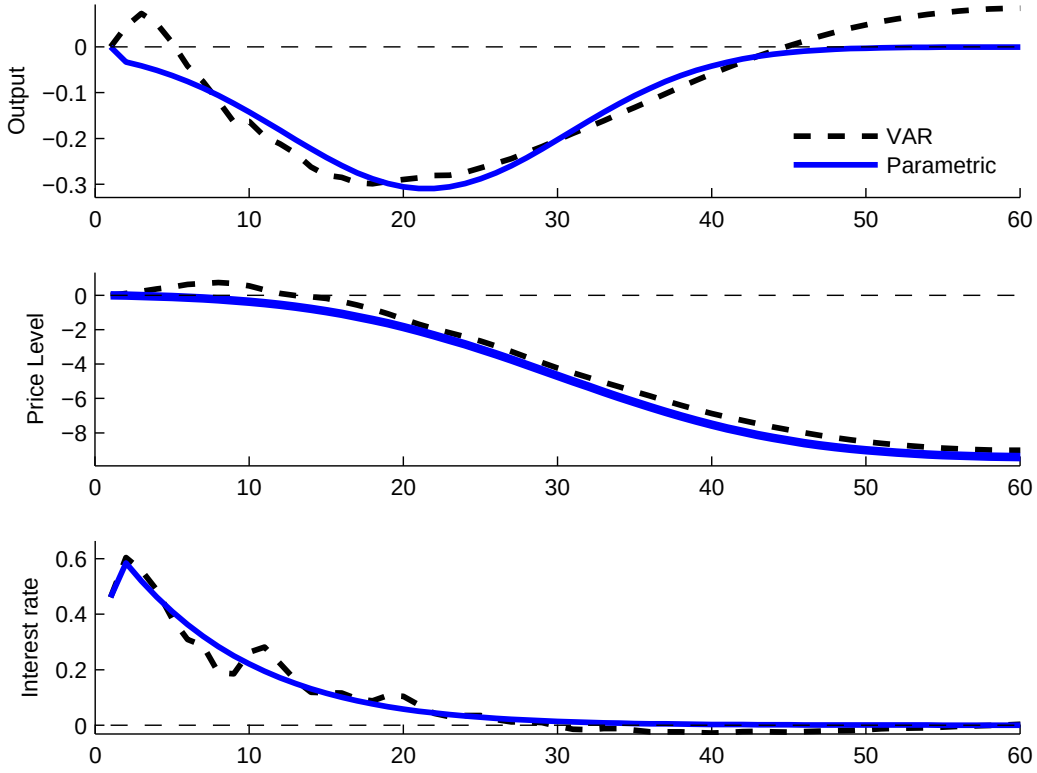


Figure 1: Impulse response functions (in percentage points) of output (industrial production), the price level and the federal funds rate to a one standard-deviation monetary shock. Estimation from a standard VAR (dashed-line) or from a parametric (one Gaussian) specification of the impulse-responses (plain line). Estimation using data covering 1959-2006.

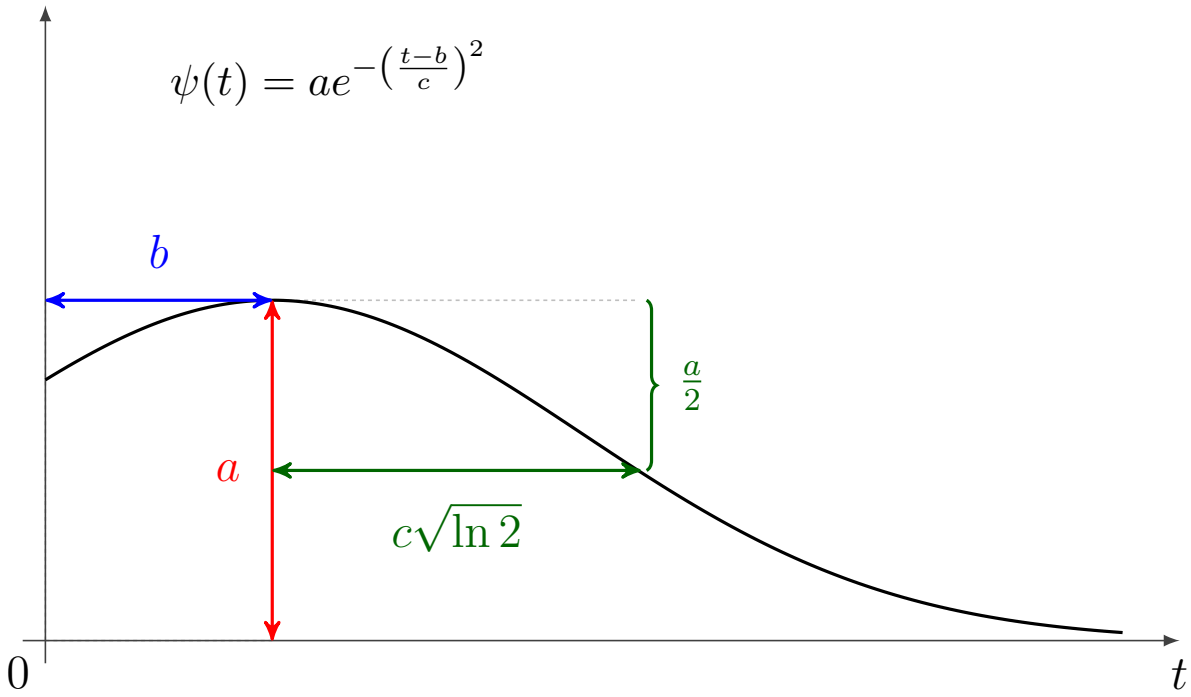


Figure 2: Interpreting an impulse response function with a Gaussian kernel.

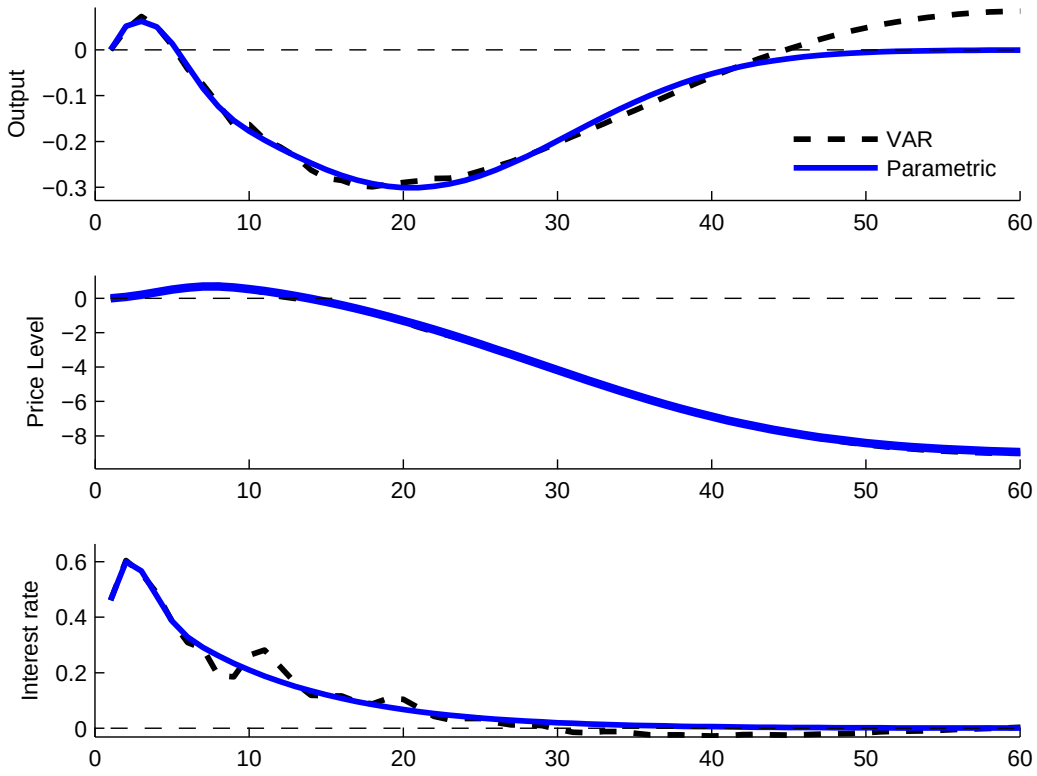


Figure 3: Impulse response functions (in percentage points) of output (industrial production), the price level and the federal funds rate to a one standard-deviation monetary shock. Estimation from a standard VAR (dashed-line) or from a parametric (two Gaussian) specification of the impulse-responses (plain line). Estimation using data covering 1959-2006.

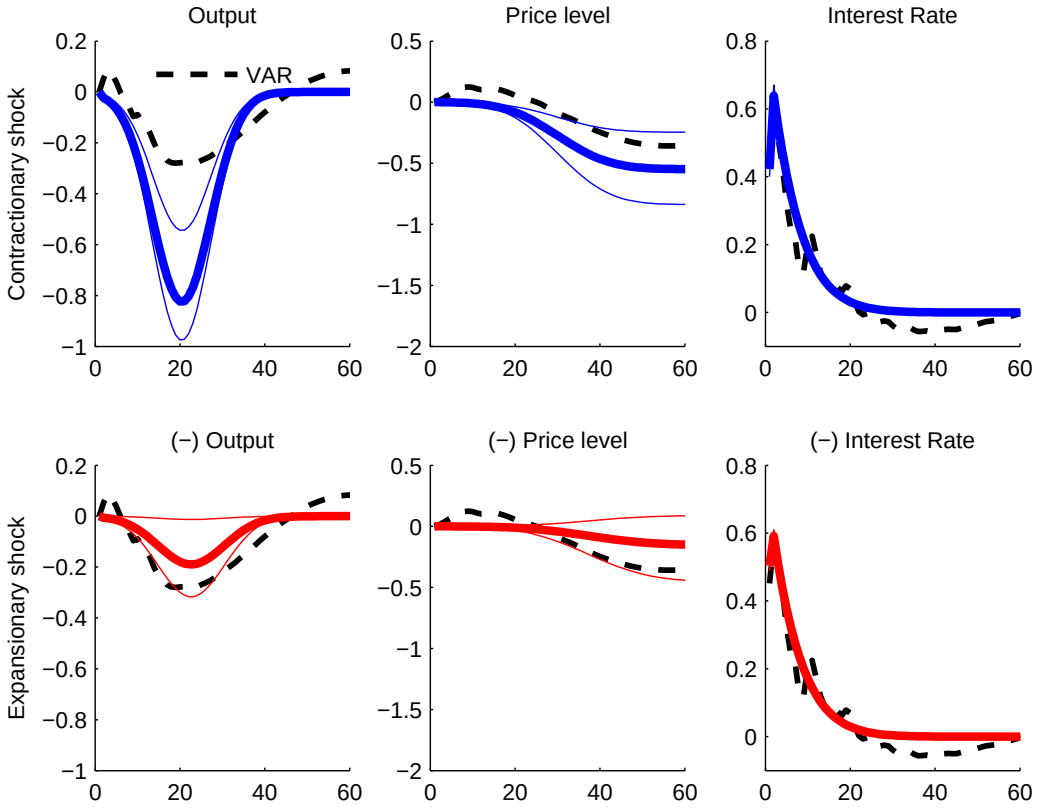


Figure 4: Impulse response functions (in percentage points) of output, the price level and the federal funds rate to a one standard-deviation monetary shock. Estimation from a standard VAR (dashed-line) or from a parametric (one-Gaussian function) specification of the impulse-responses (plain line). For ease of comparison, the effect of an expansionary shock is reported with an inverted y-axis. Estimation using data covering 1959-2006.

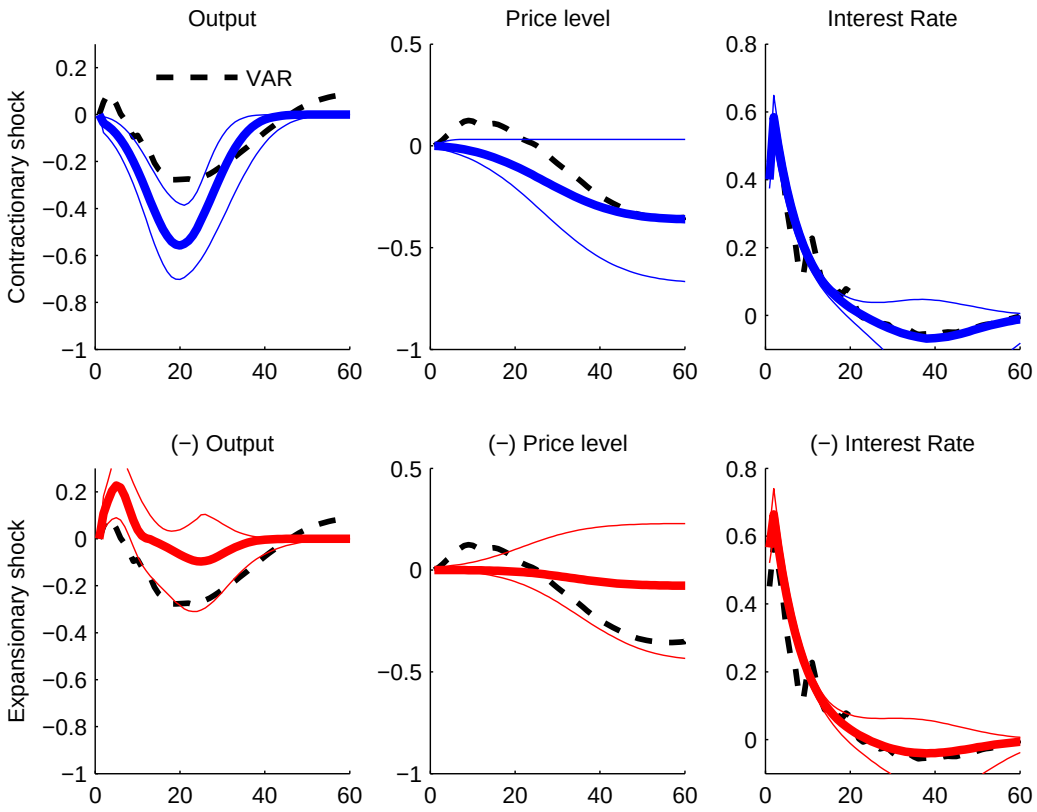


Figure 5: Impulse response functions (in percentage points) of output growth, the price level and the federal funds rate to a one standard-deviation monetary shock. Estimation from a standard VAR (dashed-line) or from a parametric (two-Gaussian functions) specification of the impulse-responses (plain line). For ease of comparison, the effect of an expansionary shock is reported with an inverted y-axis. Estimation using data covering 1959-2006.

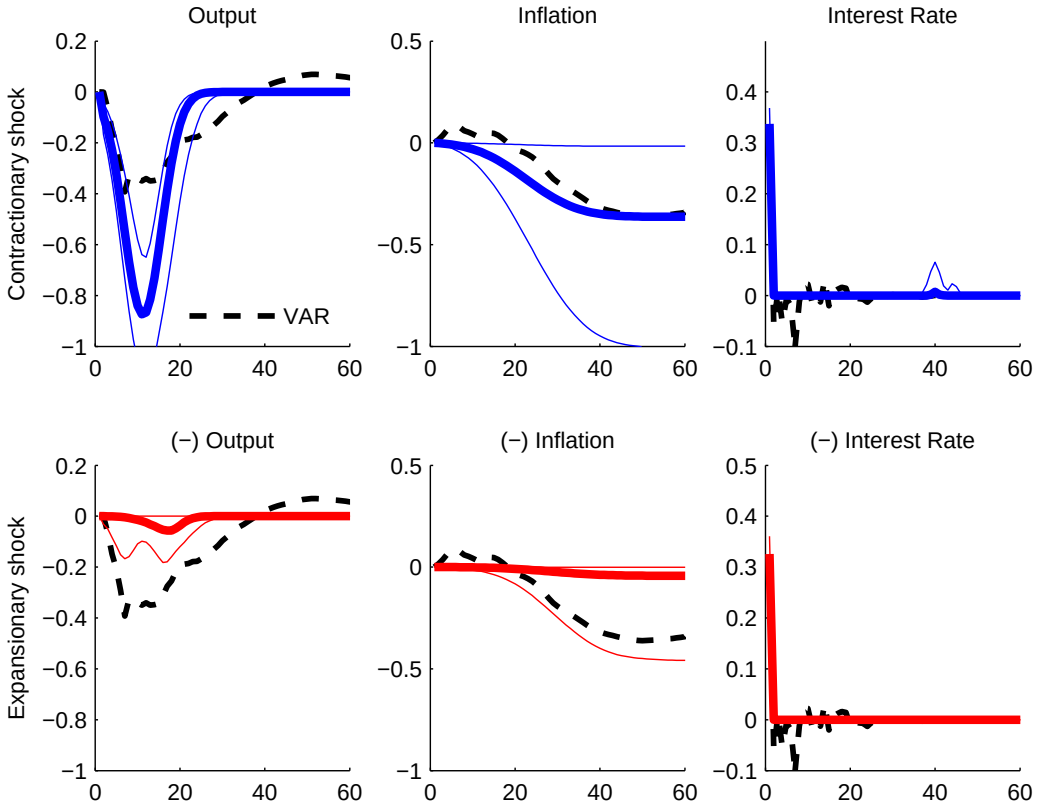


Figure 6: Impulse response functions (in percentage points) of output, the price level and the Romer and Romer monetary shock to a one standard-deviation monetary shock as estimated from a standard VAR (dashed-line) or from a parametric (one Gaussian function) specification of the impulse-responses (plain line). For ease of comparison, the effect of an expansionary shock is reported with an inverted y-axis. Estimation using data covering 1959-2006.

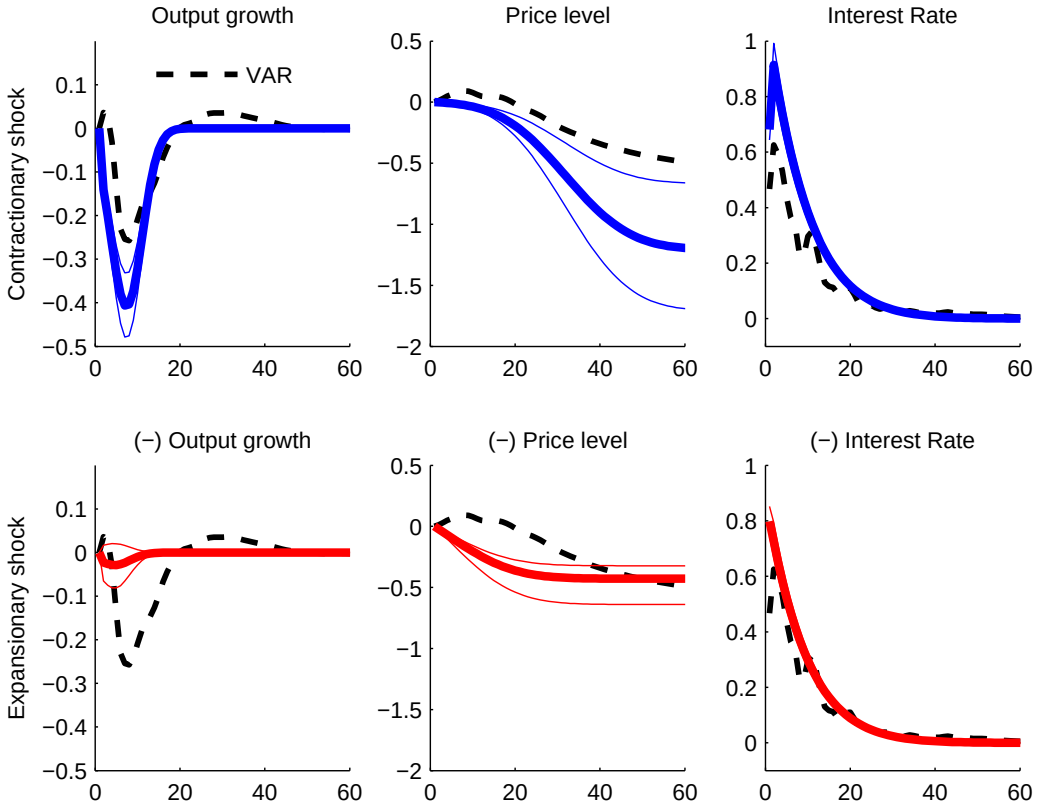


Figure 7: Impulse response functions (in percentage points) of output growth, the price level and the federal funds rate to a one standard-deviation monetary shock. Estimation from a standard VAR (dashed-line) or from a parametric (one Gaussian function) specification of the impulse-responses (plain line). For ease of comparison, the effect of an expansionary shock is reported with an inverted y-axis. Estimation using data covering 1959-2006.

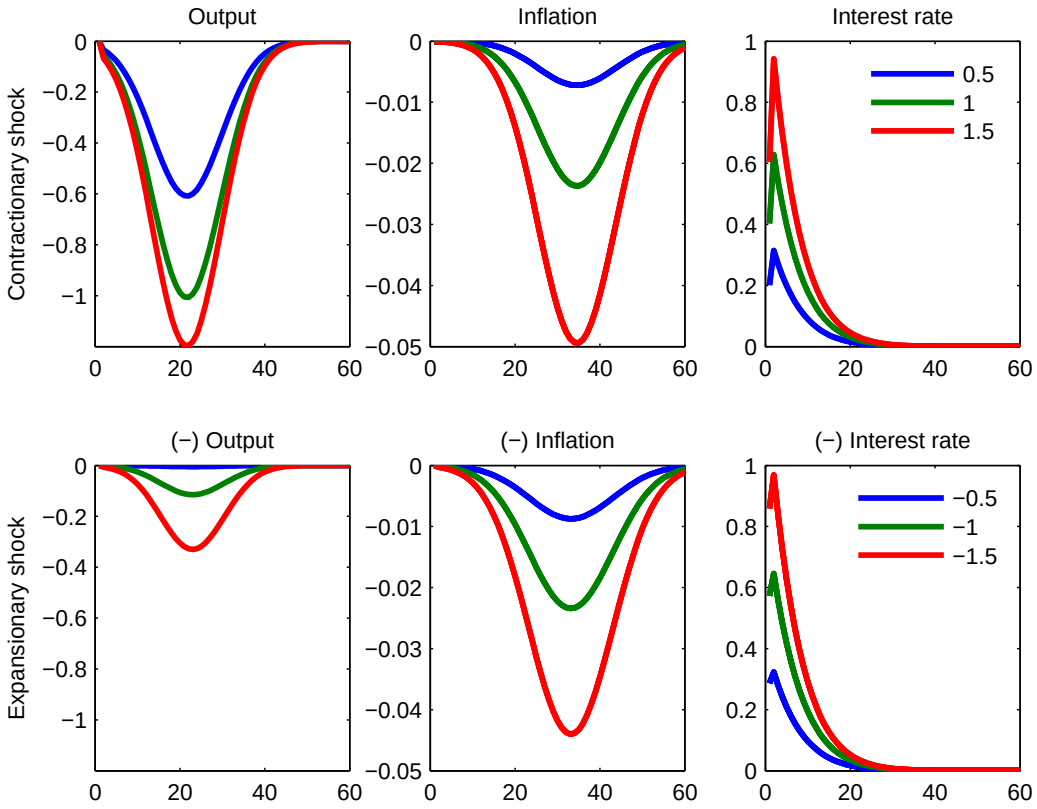


Figure 8: Impulse response functions (in percentage points) to monetary shocks of different sizes (0.5, 1 and 1.5 standard-deviations) and different signs. Estimation using a one-Gaussian parametrization with data covering 1959-2006.

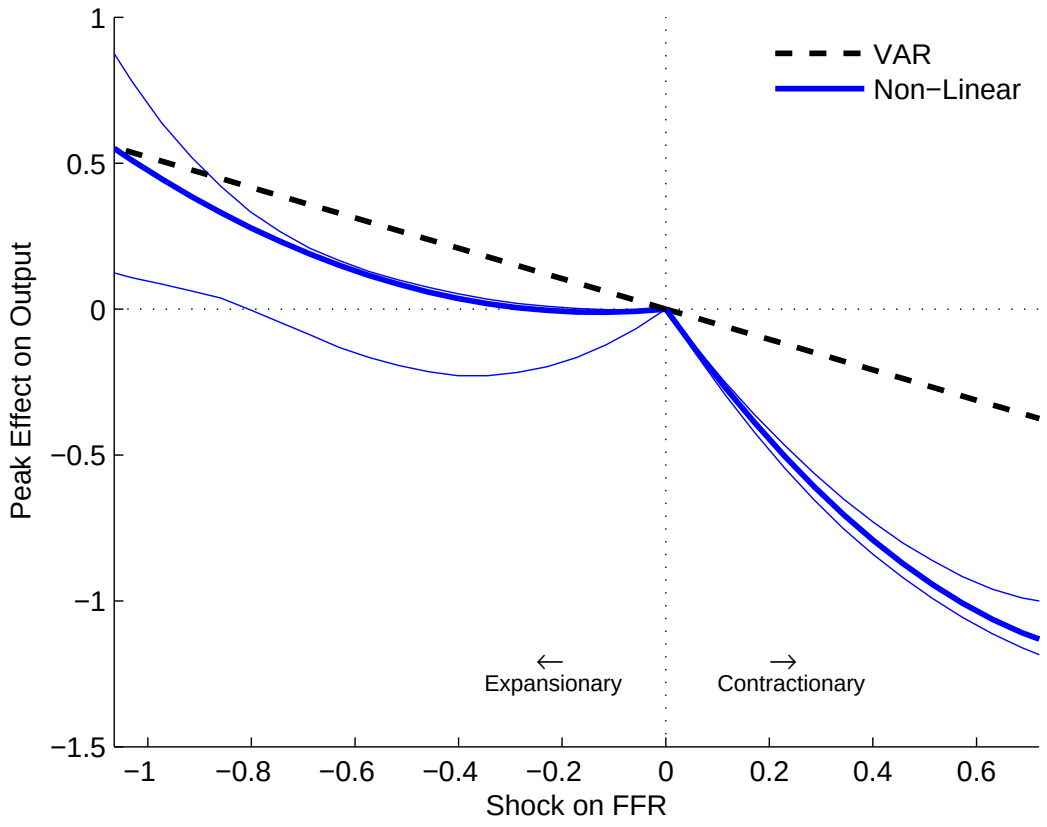


Figure 9: Peak effect of monetary policy on output (in percentage points) as a function of the size of the monetary shock (in percentage points of Fed Funds Rate, FFR). The dashed line indicate the 90% confidence interval. Estimation using data covering 1959-2006.

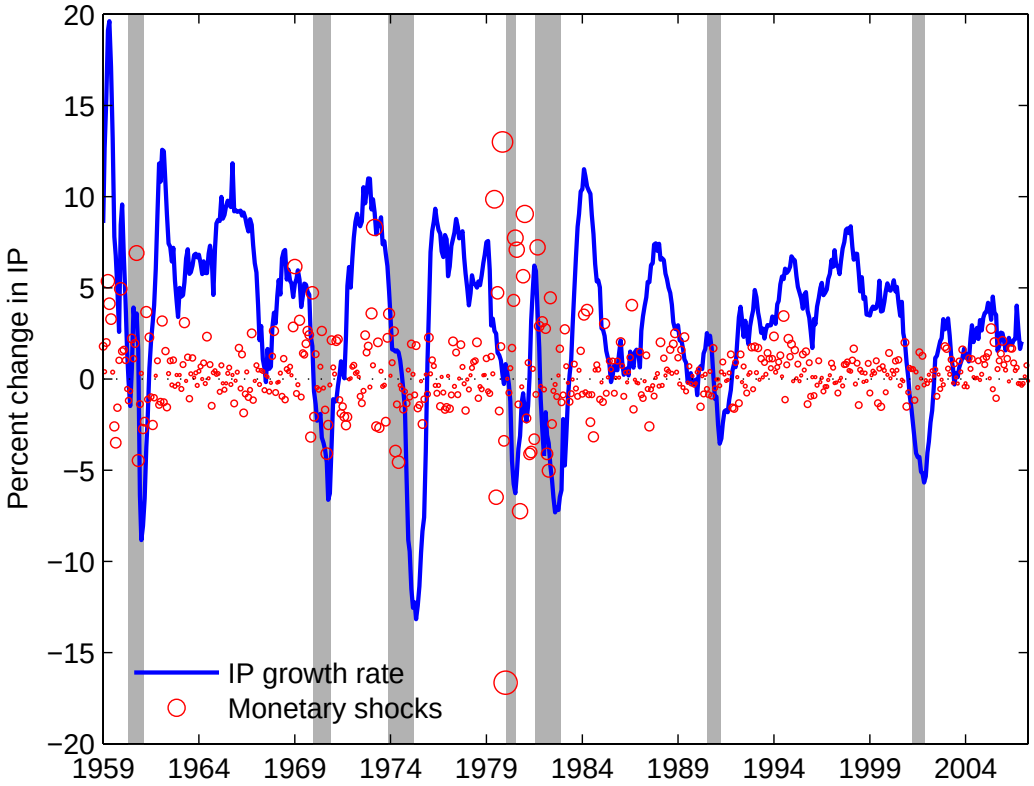


Figure 10: Annualized monthly growth rate of Industrial Production (12-months backward-moving average) used as the indicator function to capture the state of the business cycle (solid line, left scale), and estimated monetary shocks (circles, right scale) with larger circles indicating larger shocks.

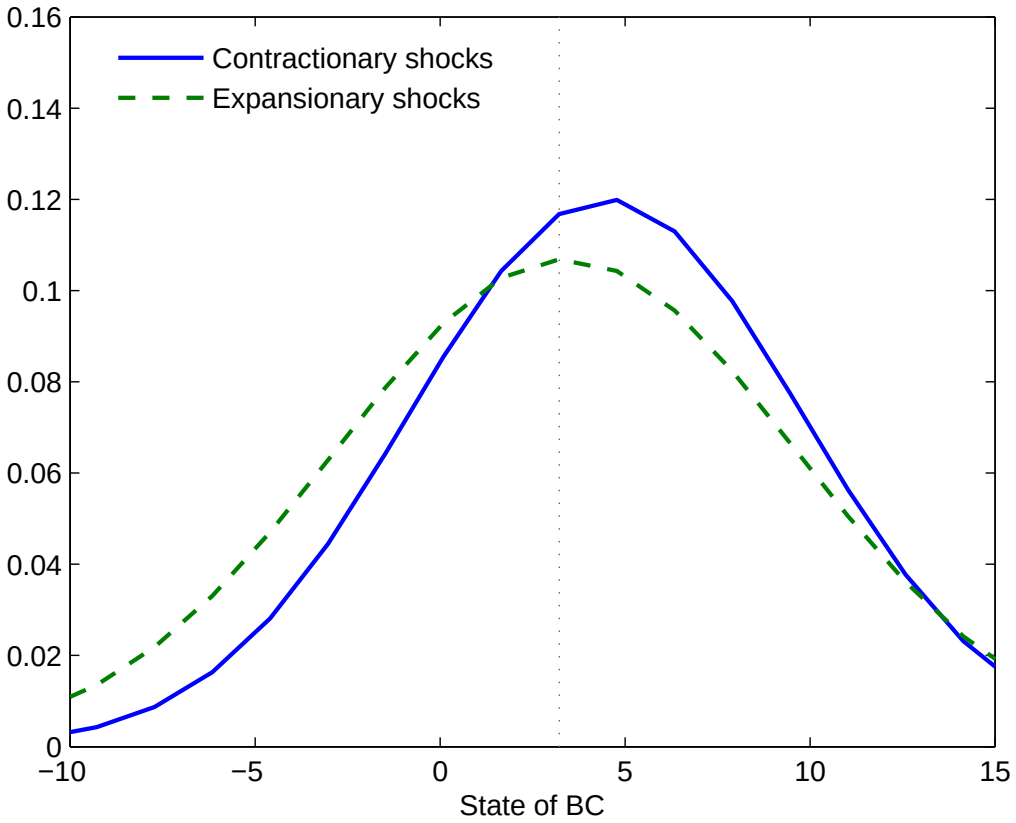


Figure 11: Distribution of expansionary shocks (dashed line) and contractionary shocks (solid line) over the state of the business cycle, defined as the annualized monthly growth rate of Industrial Production. The distributions are interpolated from a Gaussian distribution. Estimation using data covering 1959-2006.

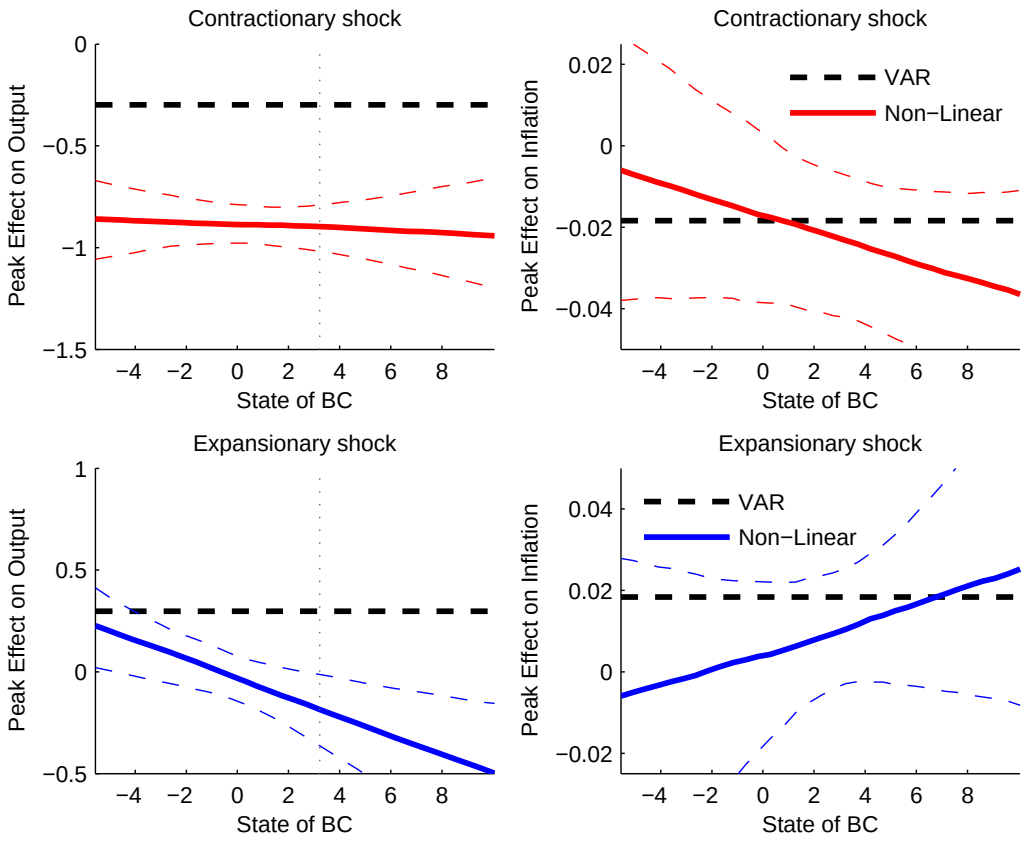


Figure 12: Peak effect of monetary policy on output and inflation (in percentage points) as a function of the state of the business cycle (in units of annualized IP growth rate) for 1 standard deviation –42 basis points– contractionary monetary shocks (upper panel) and expansionary monetary shocks (lower panel). The dashed line indicate the 90% confidence interval. The thick-dashed line is the linear VAR estimate. Estimation using data covering 1959-2006.

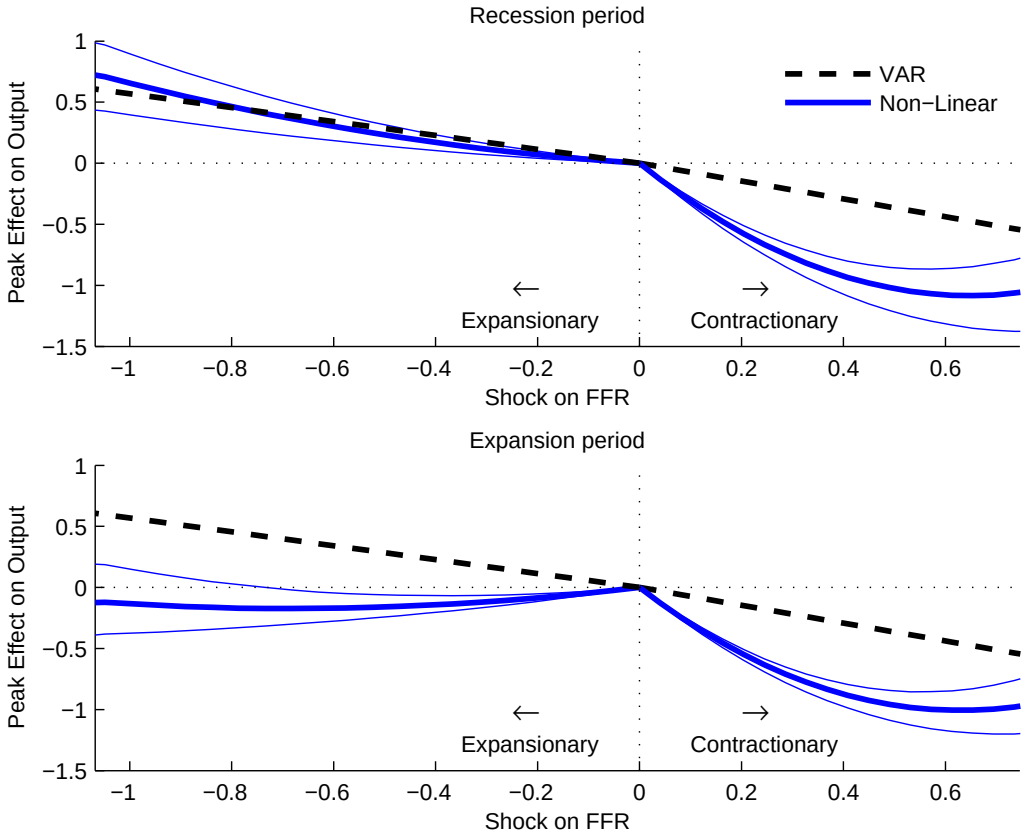


Figure 13: Peak effect of monetary policy on output (in percentage points) as a function of the size of the monetary shock (in percentage points of Fed Funds Rate, FFR) in a recession period (upper panel) and an expansionary period (lower panel). The dashed line indicate the 90% confidence interval. The thick-dashed line is the linear VAR estimate. Estimation using data covering 1959-2006.

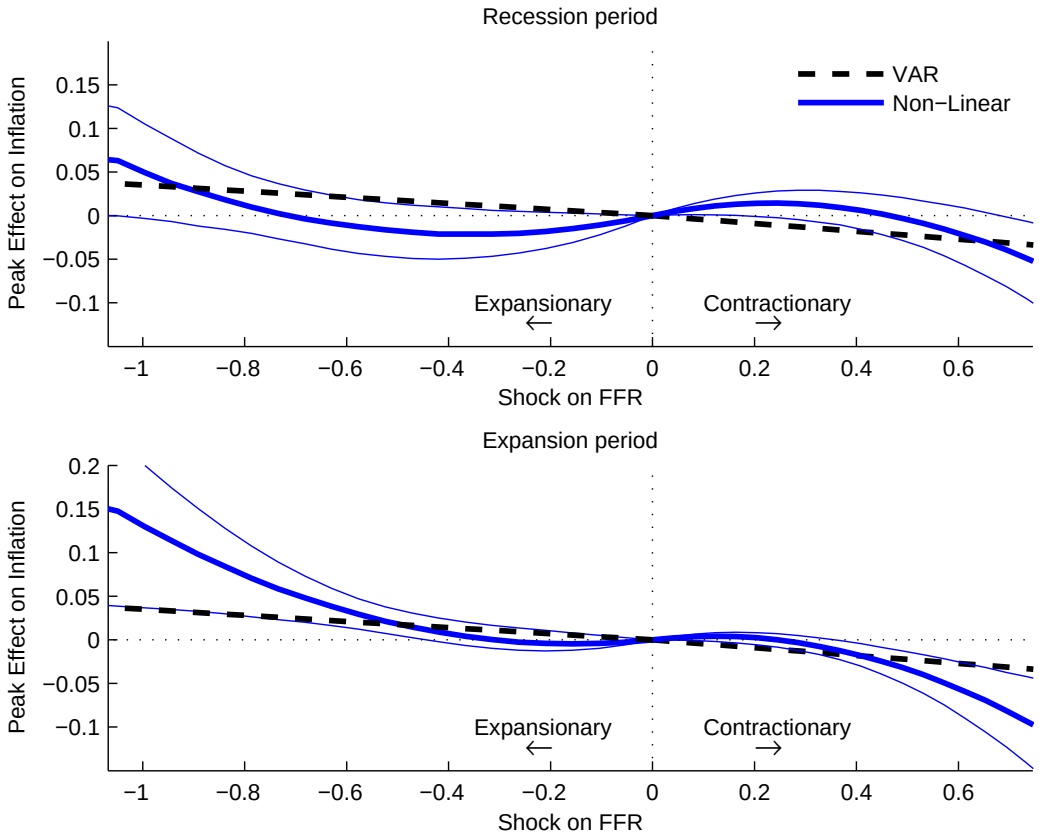


Figure 14: Peak effect of monetary policy on inflation (in percentage points) as a function of the size of the monetary shock (in percentage points of Fed Funds Rate, FFR). The dashed line indicate the 90% confidence interval. The thick-dashed line is the linear VAR estimate. Estimation using data covering 1959-2006.

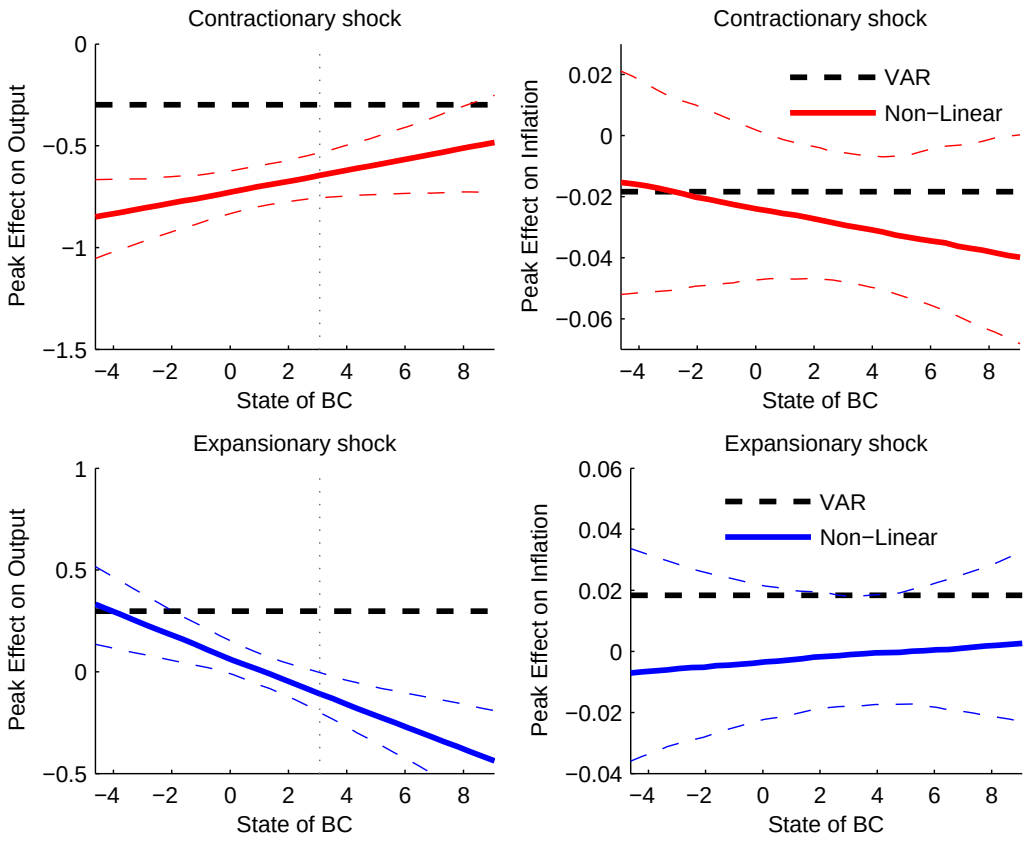


Figure 15: Peak effect of monetary policy on output and inflation (in percentage points) as a function of the state of the business cycle (in units of annualized IP growth rate) for 1 standard deviation –42 basis points– contractionary monetary shocks (upper panel) and expansionary monetary shocks (lower panel), using the Romer and Romer shocks as the policy measure. The dashed line indicate the 90% confidence interval. The thick-dashed line is the linear VAR estimate. Estimation using data covering 1966-2006.

Table 1: Log-likelihood of alternative models

	VAR	BM Symmetric One Gaussian	BM Symmetric Two Gaussian	BM Asymmetric	BM Asymmetric Size depend.	BM Asymmetric Size depend. State depend.
	(1)	(2)	(3)	(4)	(5)	(6)
(-) log-likelihood	2359	2409	2390	2247	2128	2108
Likelihood ratio test (p value)	--	--	(3) vs. (2) .90	(4) vs. (2) <.001	(5) vs. (4) <.001	(6) vs. (5) .001

Note: Specification with (log(Y),dlog(P),r) estimated over 1959m1-2006m12. Model (1) denotes the structural VAR, Model (2) is based on a parameterization of the impulse response functions (IRFs) with one Gaussian kernel, Model (3) is identical to (2) but using two Gaussian kernels and the LR test is between models (3) and (2), Model (3) is identical to (2) but allowing for asymmetric IRFs to monetary shocks and the LR test is between (4) and (2), Model (5) is identical to (4) but also allowing for the scale of the IRF to depend linearly on the size of the shock and the LR test is between models (5) and (4), Model (6) is identical to (5) but also allowing for the scale of the IRF to depend linearly on the state of the business cycle and the LR test is between models (6) and (5).

Table 2: Differences in Impulse Responses to positive and negative shocks

	Baseline 59-06	Baseline 59-06 Full asymmetry	Baseline 83-06	U and PGDP 59-06	U and PGDP 83-06
	(1)	(2)	(3)	(4)	(5)
D statistics (*100)	0.59	0.45	0.16	0.09	0.07
(95% conf. int.)	(.44, .72)	(.29, .57)	(.01, .30)	(.04, .13)	(.01, .12)

Note: Model (1) uses the variables *(log(Y),dlog(P),r)* with *Y* denoting industrial production, *P* the PCE price index and *r* the federal funds rate. (1) is estimated over 1959m1-2006m12 with a parameterization of the IRFs with a Normal kernel and allowing for asymmetric IRFs in response to monetary shocks, Model (2) is identical to (1) but allowing for asymmetric IRFs in response to all shocks, Model (3) is identical to (1) but estimated over 1983m1-2006m12, Model (3) is identical to (1) but using the Unemployment rate in level and the and the GDP deflator to measure inflation, Model (4) is identical to (5) but estimated over 1983m1-2006m12.