INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM

WORKING PAPER SERIES

Macroeconomic responses to an independent monetary policy shock: a (more) agnostic identification procedure

Marco Capasso °
Alessio Moneta °

° Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy

2016/36 October 2016
ISSN(ONLINE) 2284-0400

Macroeconomic responses to an independent monetary policy shock: a (more) agnostic identification procedure

Marco Capasso* Alessio Moneta†

October 21, 2016

Abstract

This study investigates the effects of a monetary policy shock on real output and prices, by means of a novel distribution-free nonrecursive identification scheme for structural vector autoregressions. Structural shocks are assumed to be mutually independent. The identification procedure is agnostic in Uhlig [2005]'s sense, since the response of output to a monetary shock is not restricted. Moreover, assuming mutual independence of the shocks allows us to impose no additional constraints derived from economic theory.

Keywords: Structural Models, Vector Autoregressions, Independent Component Analysis, Identification, Monetary Policy.

JEL Classification: C14 · C18 · C38 · C51 · E52.

*Scuola Superiore Sant'Anna, Pisa, Italy, and Finance Think, Skopje, Macedonia.

†Corresponding author. Scuola Superiore Sant'Anna, Piazza Martiri della Libertà 33, 56127 Pisa, Italy. E-mail: a.moneta@sssup.it.

This work was funded by the Institute for New Economic Thinking under grant agreement INO15-00021.

We thank Giulio Bottazzi, Daniele Giachini, Irene Iodice and all the participants of the Institute of Economics Seminar at Scuola Superiore Sant'Anna, July 2016, for useful comments.

1 Economic Motivation

Our study is aimed at answering two questions. Does a contractionary monetary policy reduce the real gross domestic product? Does a contractionary monetary policy reduce prices?
We answer the two questions above within the framework of Vector AutoRegressions (VARs). Sims [1980] introduces VARs by loosening some of the theoretical economic assumptions constraining the simultaneous equation models, and suggests a first "structural" version of vector autoregressions to show that, in the United States, an exogenous increase in the money stock translates into a temporary increase of the real gross national product and into a permanent increase in prices (for a one standard deviation shock, prices increase for about two years). A deeper reflection on the coexistence of supply and demand elements in the money stock brings Sims [1986] to argue that, after an unpredicted increase in Treasury-Bill rates, the real gross national product would decrease, while the response of prices would first be positive (up to around three years) and then negative.

The latter result is stylized by Sims [1992] [cfr. also Eichenbaum, 1992] as the "price puzzle": positive interest rate innovations, interpretable as contractionary monetary policy shocks, seem to generate an increase in the general level of prices. "Policy authorities might know that inflationary pressure is about to arrive", Sims [1992, p. 988] suggests, and the price puzzle could thus result from an endogeneity in monetary policy which escapes the chronological order suggested by the observable data. On the other hand, Sims [1992] shows that commodity prices immediately decrease following the contractionary policy (one of the reasons why the subsequent studies have often included commodity prices in the macroeconomic datasets, see also the comments by Christiano et al., 1999, on this issue).
Starting with the article by Bernanke and Blinder [1992], exogenous shocks to the Federal funds rate are labelled as unanticipated changes in monetary policy, or monetary policy shocks. Using monthly data, Bernanke and Gertler [1995] and Bernanke and Mihov [1998] confirm the results of the previous literature, although with a changed time frame: following a contractionary monetary policy (an exogenous increase in the Federal funds rate), the response of the real gross domestic product (GDP) becomes negative within four months, while the response of the GDP deflator becomes negative after about one year of being positive.

Uhlig [2005] argues that the existing literature, including the works we have mentioned above, adopts identifying assumptions which look reasonable *a priori*, based on economic theory, but less reasonable *ex post*, because they lead to puzzling results. Moreover, the attempts made in the literature to attain more credible results, by adopting different economic assumptions, often narrow down the range of possible answers to the question of interest, about the impact of monetary policy shocks on GDP, by mechanisms which are not always explicit. As a complement to the existing literature, it would thus be desirable "to make the a priori theorizing explicit (and use as little of it as possible), while at the same time leaving the question of interest open" [Uhlig, 2005, p. 384]. A new identification procedure is then suggested, under the assumption that a contractionary monetary policy shock does not lead to increases in prices, nor to increases in nonborrowed reserves, nor to decreases in the federal funds rate, for a certain period following the shock. The procedure is called "agnostic" because the key question, about the effects of monetary policy on output, is left "agnostically open by design of the identification procedure: the data will decide" [Uhlig, 2005, p. 384]. Results show that, following a contractionary monetary policy shock, prices fall (slowly in terms of the GDP price deflator, quickly in terms of the commodity price index). The real effects are ambiguous, with a higher probability of the real GDP increasing instead of decreasing.

Arias et al. [2015] have recently criticized the restrictions imposed by Uhlig [2005], while endorsing his agnostic identification approach. Instead of constraining the responses to the monetary policy exogenous shock, they suggest to identify the structural shocks by imposing restrictions on the systematic (endogenous) component of monetary policy. Motivated by Christiano et al. [1996], their baseline identification scheme assumes that the Federal Funds rate only reacts contemporaneously to output and prices, and the contemporaneous reaction is nonnegative (other identification schemes are also proposed, still based on zero and sign restrictions on the systematic component of monetary policy). Results show that a contractionary monetary policy leads to a persistent decline in output and prices.

We follow the agnostic approach of Uhlig [2005]: we want to make theoretical assumptions whose constraint on the answer to the key question, about the impact of monetary policy shocks on GDP, is as small as possible, and as explicit as possible. Like Arias et al. [2015], we want our theoretical assumptions not to exclude positive effects of contractionary monetary policies on prices. Our intuition for identification is: all the works cited above assume orthogonality of structural shocks, but the economic reasons behind orthogonality would also support mutual independence of the structural shocks. Since independence is a stronger assumption than orthogonality, identification of the structural vector autoregression may be freed from other theory-based constraints. Moneta et al. [2013] have recently shown the possibilities offered by independent component analysis, and in particular by the algorithms in Hyvärinen et al. [2010], for analyzing the effects of monetary policy. We suggest a simplification of their identification procedure, and apply it on the same data used by Uhlig [2005] and Arias et al. [2015].

2 Model

Our model is defined by the following four assumptions:

(Assumption 1) Each observable variable is a random variable, whose realization at each time period depends only and linearly on the previous realizations of the same variable, on the previous realizations of all the other observable variables, and on unpredictable mutually independent random shocks which we call "structural shocks" (each observable variable may be hit by the structural shocks directly, or contemporaneously through other observable variables);

(Assumption 2) The coefficients describing the linear dependencies, among the observable variables as well as between the observable variables and the unpredictable random shocks, are constant over time;

(Assumption 3) At each time period, the same number of structural shocks impact the observable variables; each of the structural shocks is drawn according to a probability distribution which is constant over time; each drawing is independent from previous drawings;

(Assumption 4) No more than one structural shock is drawn from a Gaussian probability distribution;

(Assumption 5) The number of structural shocks is equal to the number of observable variables.

The first part of Assumption 1 and Assumption 2 place our model within the standard framework of vector autoregressions: the current realization of any observed variable may influence the future realization of all variables. In notation:

$$x_t = D_1 x_{t-1} + D_2 x_{t-2} + \cdots + D_h x_{t-h} + u_t \ , \tag{1}$$

where x_t is the vector of realizations at time t of the observable variables; h is the maximum lag of the vector autoregression; $D_1, D_2 \ldots D_h$ are the parameters describing the relations between past and current values of the observable variables; u_t is the vector of residuals at time t i.e. the difference between the observed x_t and the value of x_t predicted at time $t-1$.
The second part of Assumption 1 implies that the residuals in (1) are linear combinations of mutually independent random shocks:

$$u_t = A\varepsilon_t \ , \tag{2}$$

where ε_t is the vector containing the realizations at time t of the random mutually independent shocks.

The mutually independent shocks ε are called "structural", in the broad sense of "connected to a claim of usefulness in the prediction of the effects of actions" (Sims, 1986, p. 4, footnote 2). Indeed, by reconducting the reduced-form residuals to a set of independent shocks, we want to disentangle the independent sources of variation in the system. Since the introduction of vector autoregressions, residual orthogonalization has been a way to see "the distinct patterns of movement the system may display" (Sims, 1980, p. 21). We think that the arguments usually put forward to justify the orthogonality of structural shocks are, in many contexts, consistent with the stronger assumption of independence of structural shocks. We will use this stronger assumption for identification.

Assumption 3 states that the structural shocks are serially independent and identically distributed, and allows us to infer more easily the shock probability distributions from the observable data.
Assumption 4 is made only in order to apply Independent Component Analysis, which will be described in the next section; Assumption 5 is made only for simplicity and to facilitate comparisons with other autoregressive models. We do not have a motivation rooted in economic theory for choosing a particular number of random shocks driving the economic system, nor are there statistical constraints which would otherwise prevent our model estimation (in the last paragraph of Section 4, we give a hint on how Assumption 5 could be relaxed).

3 Estimation

First, each observable variable is regressed, by ordinary least squares, on its own past values and on the past values of all the other observable variables (one separate regression for each observed variable). In this way, we obtain an estimate of the D matrices in (1), and, for each time t, of the residuals u_t. The estimates $\hat{u}_t$ are the differences between observed and predicted values of x_t. Roughly speaking, the regressions are used to wash away the effect of the past on x_t, so that our analysis may focus on the innovations of x_t.
Independent component analysis is then applied on the matrix $\hat{U}$ having as columns all the vectors $\hat{u}_t$ obtained as residuals for each time t. Indeed, following the notation of (2), Assumption 1 implies that, for any t, $\hat{u}_t$ derives from a vector of independent "structural" shocks $\hat{\varepsilon}_t$, linearly combined according to a mixing matrix A which is constant over time. If we denote by $\mathcal{E}$ the matrix having as columns all the structural shocks $\hat{\varepsilon}_t$ which impact the system at each time t, then the independent component analysis can be defined as the search for the unmixing matrix W which, multiplied by U, results in rows of $\mathcal{E}$ "as independent as possible":

$$WU = \mathcal{E} \ , \tag{3}$$

(W is the inverse of the mixing matrix A defined in the previous section).

Following Comon [1994], we define "as independent as possible" the structural shocks for which the distribution of the product of marginal probabilities diverges the least from the joint distribution, where divergence is measured by the "mean information for discrimination" (Kullback and Leibler, 1951, p. 80). In our context, assuming that the system dynamics is driven by n mutually independent shocks (n being also the number of observable variables), the goal of our structural analysis is finding the n linear combinations $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ of the reduced-form residuals $u_1, u_2, \dots, u_n$ such that the following Kullback-Leibler divergence is minimized:

$$\int_{\mathcal{E}_n} \cdots \int_{\mathcal{E}_2} \int_{\mathcal{E}_1} p(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) \log \frac{p(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)}{p(\varepsilon_1)p(\varepsilon_2)\dots p(\varepsilon_n)} \mathrm{d}\varepsilon_1 \, \mathrm{d}\varepsilon_2 \, \dots \, d\varepsilon_n \ , \tag{4}$$

where $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_n$ indicate the support of the marginal distributions of respectively $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$.
The Kullback-Leibler divergence of the distribution of the product of marginal probabilities from the joint probability distribution, represented in (4), is often called "mutual information" (see, e.g., Fraser and Swinney, 1986, p. 1137). Our

search for minimum dependence among structural shocks can thus be seen as a search for shocks having minimum mutual information.[1]

Comon [1994, pp. 295-296] has shown that the divergence in (4) can be minimized by maximizing the negentropy of the marginal distributions of respectively $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, under the assumptions that $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ are uncorrelated and at most one of them is Gaussian. Negentropy is defined as the result of a particular normalization of differential entropy: while the differential entropy of a random shock ε_i is defined as

$$H(\varepsilon_i) = -\int_{\mathcal{E}_i} p(\varepsilon_i) \log p(\varepsilon_i) d\varepsilon_i \ , \tag{5}$$

the negentropy of ε_i is defined as the opposite of the difference between $H(\varepsilon_i)$ and the differential entropy of a Gaussian random variable having same variance as ε_i. The negentropy of a probability distribution can thus be interpreted as a distance from Gaussianity.

Hyvärinen [1999] has suggested an efficient algorithm, named "FastICA", to find the directions in which negentropy is maximized. Given that the estimates of the FastICA algorithm may change according to its initialization, Himberg and Hyvärinen [2003] have built the "Icasso" software package to obtain more robust estimates, by averaging FastICA results obtained under different initializations. The Icasso package has a consolidated reputation within the neuroscience research community (see e.g. Himberg et al., 2004, and Corradi-Dell'Acqua et al., 2016). We will feed Icasso with the residuals $u_1, u_2, \ldots, u_n$ of the reduced-form VAR estimation, in order to retrieve the structural shocks $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$.[2]

[1] For an overview of the possible contributions of information theory to econometrics, see the special issue introduced by Golan [2002].

[2] We use: version 1.22 of Icasso,
downloaded from `http://research.ics.aalto.fi/ica/icasso/`;
version 2.5 of FastICA,
downloaded from `http://research.ics.aalto.fi/ica/fastica/`;
release R2015b of MATLAB,
institutionally licensed (`http://www.mathworks.com/products/matlab/`).
The additional MATLAB code written for this study is available at `www.capasso.info`.

4 Methodological Novelty

Moneta et al. [2013] have recently highlighted the possibilities offered by Independent Component Analysis in both micro- and macro-economic contexts. In particular, they have shown that the approach by Hyvärinen et al. [2010], who model the residuals of a reduced-form VAR estimation according to the "Linear Non-Gaussian Acyclic Model" (LiNGAM) by Shimizu et al. [2006], can be fruitful when there is no consensus on the economic theories to use for identification. While our study also applies independent component analysis to VAR residuals, there is one main feature in the LiNGAM approach that we want to jettison. The LiNGAM would assume a recursive contemporaneous causal structure among the observed variables ("acyclicality" assumption), thus forcing the unmixing matrix, named W in our equation 3, to be triangular (see Algorithm C at page 2009 of Shimizu et al., 2006, and step 6 of the algorithm at page 715 of Moneta et al., 2013).

In the previous literature, recursiveness has been a usual assumption when structural VARs are identified by means of short-run restrictions. Starting with Sims [1980, p. 21], economists have often imposed particular acyclic schemes to the contemporaneous relations among observed variables, chosen on the basis of theoretical economic reflections. Stock and Watson [2001, p. 112] have warned that "[r]arely does it add value to repackage a recursive VAR and sell it as structural", and that a deeper inspection on previous VAR analyses often reveals flaws in the economic theories behind particular recursive schemes, or behind recursiveness in general. The graph-theoretic approach to VAR identification proposed by Bessler and Lee [2002], Demiralp and Hoover [2003], and Moneta [2008], like the LiNGAM approach by Hyvärinen et al. [2010] and Moneta et al. [2013], avoids the *a priori* choice of a particular recursive scheme, and suggests a procedure to select the most plausible causal structure among the contemporaneous variables according to the data, where the contemporaneous causal structure is assumed to be represented as a directed acyclic graph. This is equivalent to say that there exists a recursive causal scheme (of which we do not know the order) that constitutes the propagation mechanism of the shocks. Moreover, the structural shocks are associated one-to-one to the observable variables: at each time period, each structural shock is assumed to impact one particular observable variable, and instantaneously propagate throughout the system because of the causal relations among the observable variables.

Instead, we are closer in spirit to the view of Blanchard and Quah [1989], and in general of VAR modellers who consider long-run identification schemes. Blanchard and Quah [1989] assume that two unobservable random shocks, respectively "demand" and "supply", drive the dynamics of the observable variables "output" and "unemployment". Both demand and supply shocks affect contemporaneously both output and unemployment, and it would be difficult, at least at yearly data frequency, to assume a recursive scheme in which one of the unobservable shocks only impacts one of the two observable variables but not the other one. No structural shock is considered as an exogenous innovation of one particular observed variable. In our paper, although we do not impose long-run restrictions, structural shocks do not necessarily have a one-to-one association to observed variables. They might even be latent factors affecting contemporaneously different observable variables that would otherwise be causally disconnected. Recursiveness is allowed as a possibility, not imposed as a restriction.

We take stock of the remarks by Lanne et al. [forthcoming], who use structural shock independence as an identification restriction, and strongly argue against assuming recursiveness. However, we do not follow them on the parametric estimation of the mutually independent structural shocks, since we prefer not to assume any particular non-Gaussian probability distribution (or family of distributions) for the structural shocks. In their macroeconomic application, Lanne et al. [forthcoming] impose that the structural shocks are drawn from Student's t-distributions. Lanne and Lütkepohl [2010] have previously suggested to assume normal mixture distributions of the structural shocks, when the macroeconomic system is characterized by two different regimes, with different volatilities. Siegfried [2002] has argued that monetary policy shocks should be distributed logistically, based on an analysis of the decision-making processes within central banks.
We think that, in our empirical work, there is no solid ground on which the assumption of a particular shock distribution could rest. Although Gouriéroux et al. [forthcoming] have recently shown that a misspecification of the prior distributions of the independent shocks does not lead to inconsistent estimates, we follow a distribution-free approach similar to the one proposed by Herwartz and Plödt [2016], who assume neither recursiveness nor a particular probability distribution of the structural shocks. However, Herwartz and Plödt [2016] search for the minimal distance between the empirical copula of the shock vector and the theoretical copula under the hypothesis of independence and, consequently, they use a different estimation procedure than ours.

A possibility that we leave open for future analyses is relaxing Assumption 5. For instance, we could imagine an approximate factor model (as in Chamberlain and Rothschild, 1983), where the dependencies among the n observed variables are mainly due to the common influence of q mutually independent random shocks (with $q < n$). Our estimation procedure could easily be adapted by considering only the first q principal components of our dataset, and rotating only those q principal components to retrieve the q mutually independent structural shocks. The data dependencies left unexplained by the q structural shocks would be ascribed to spurious dependencies among idiosyncratic variations (for instance, measurement errors) of the observed variables.

5 Empirical Analysis

We use the dataset in Uhlig [2005], which contains macroeconomic observations for the United States since January 1965 until December 2003 (a similar dataset had previously been employed, with a shorter time span, by Bernanke and Mihov, 1998). The six observable variables are: real gross domestic product (GDP), GDP price deflator (PGDP), commodity price index (PSCCOM), total reserves (TR), non-borrowed reserves (NBR), Federal funds rate (FFR).[3] All data have a monthly frequency following the interpolation of GDP and GDP deflator; details about the dataset construction are in Section 3 of Uhlig, 2005. (We also perform estimations on quarterly data using the variables in Sims, 1986; results are shown in Appendix A). We follow Uhlig [2005] in the choice of taking the logarithm of all the variables except for FFR, and in the number of lags, equal to 12, of the VAR. We thus estimate, by ordinary least squares, an unrestricted 12-lag VAR (in levels) in the reduced form of (1). Figure 1 contains the histograms of the reduced-form residuals, associated to each of the six observed variables; a normal distribution fit has been added to highlight their leptokurtosis.[4]

We use Icasso to estimate (2), that is to retrieve the mixing matrix and the structural

[3] The dataset is freely available in the compressed folder:
`https://estima.com/procs_perl/uhligjme2005.zip`.
In particular, we have used the file `uhligdata.xls`, last modified on February 23, 2006.

[4] The histograms have been obtained using the "histfit" MATLAB function. The estimates of the reduced-form parameters are available upon request.

	ε_1	ε_2	ε_3	ε_4	ε_5	ε_6
GDP	0.016	0.000	-0.050	-0.298	0.066	0.023
PGDP	-0.018	-0.004	-0.010	-0.015	-0.092	0.056
PSCCOM	0.142	0.448	0.323	0.369	1.051	2.466
TR	-2.148	0.089	-0.145	0.063	0.025	0.038
NBR	-1.799	1.339	0.410	0.031	0.019	-0.221
FFR	0.012	0.018	-0.485	0.054	0.041	0.061

Table 1: Mixing matrix estimate. Each row reports how the six shocks load on the reduced-form residual associated to one particular observed variable.

independent shocks from the reduced-form residuals.[5] Given that the structural independent shocks estimated by Icasso are not precisely orthonormal, we rescale them so that they have variances equal to one, and we accordingly rescale the mixing matrix. The resulting mixing matrix, i.e. the estimate of the matrix A in (2), is shown in Table 1: each row reports how the six shocks load on the reduced-form residual associated to one particular observed variable (rows are ordered following the variable order given at the beginning of this section).

We label as "monetary policy" the structural shock which contributes the most to the FFR's variance. The sixth row of the mixing matrix in Table 1 shows how the six structural shocks load on the reduced-form residual associated to FFR. The highest number in the sixth row, in absolute value, occupies the third cell. Therefore, we label as "monetary policy" the structural shock associated to the third column of the mixing matrix (different labelling possibilities are explored in Appendix B). To obtain the impulse-response functions, we generate a contractionary "monetary policy" impulse in the form of a vector where all cells contain zeroes except for the loading on FFR which is equal to negative one (given the rescaling, this corresponds to a one standard deviation monetary policy impulse, raising FFR at impact). Then, using the mixing matrix in Table 1, we obtain the response at lag zero (i.e. the impact of the impulse on the observed variables), and using the estimates of the reduced form in (1) we obtain the responses for lags greater than zero.

[5] We let Icasso run 500 iterations of FastICA, with a different initialization for each iteration. For each given initialization, we allow FastICA a maximum number of 300 steps to reach convergence; if convergence is not reached, the value obtained by the algorithm after 300 steps is kept for the Icasso averaging.

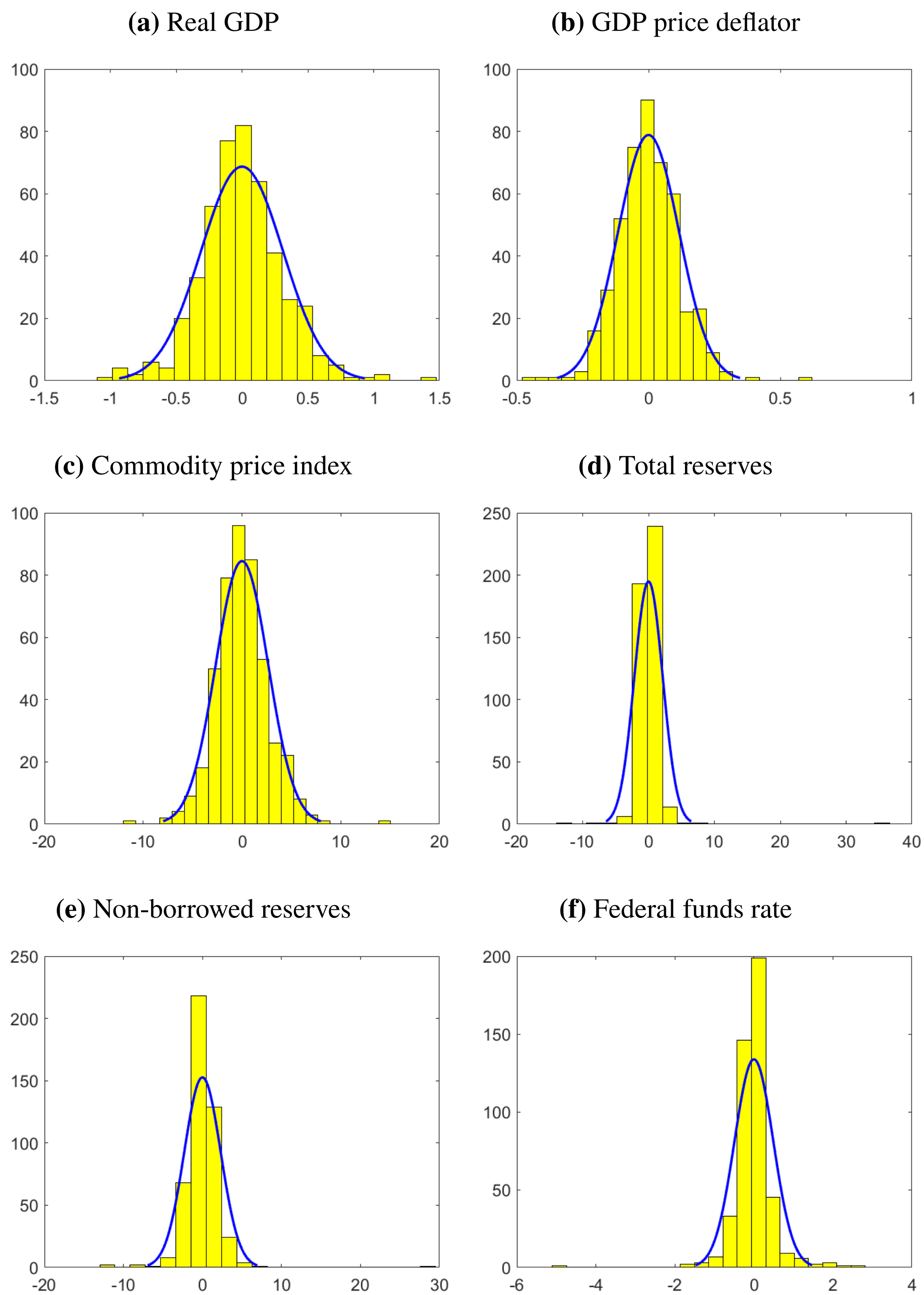

Figure 1: Histograms of the reduced-form residuals, with a normal distribution fit.

In order to obtain confidence intervals for the impulse-response functions, a bootstrap procedure is used consisting of 1000 iterations. At each iteration, a matrix of simulated residuals $\tilde{U}$ is generated on the basis of the reduced-form residuals in $\hat{U}$ (as defined in Section 3). In particular, for $k = 1, \ldots, n$, the k-th row of $\tilde{U}$, with length equal to the number of observations T in the dataset minus the number of VAR lags h, is obtained by drawing, with repetition, $T - h$ values from a discrete probability distribution which is uniform on the values contained in the k-th row of $\hat{U}$ (i.e. which is uniform on the estimates of the reduced-form residuals). The simulated residuals in $\tilde{U}$ are fed to Icasso, so that a new mixing matrix and the corresponding impulse-response functions can be obtained and stored. At the end of all the 1000 iterations, the quantiles of the distribution of responses, computed separately for each response lag, are used to build the confidence intervals of the impulse-response functions.

The resulting impulse-response functions, estimated as responses of each of the six variables to the contractionary monetary policy impulse, are shown in Figure 2. Immediately after a contractionary monetary policy shock, commodity prices fall. After around six months, the real GDP starts to decline, and the response remains negative thereafter. The response of the GDP price deflator is initially positive, but becomes negative after about three years.

Our findings are thus far from what shown in Uhlig [2005]: both the GDP contraction and the price puzzle are evident. The panels of Figure 2 remind of the impulse responses obtained in many previous studies, like in Sims [1986], in Sims [1992], in Christiano et al. [1996], and in Bernanke and Mihov [1998]. In particular, Figure 2 bears a striking resemblance to Figures 7 and 12 in Arias et al. [2015], where, while estimating the model on the same data as ours, identification has been achieved, respectively, by adopting the recursive scheme in Christiano et al. [1996] and by assuming a systematic policy response to changes in the M2 money stock. Notably, we reach their same result by imposing shock independence as the main identifying assumption.

6 Conclusion

Our study has applied a novel identification scheme to a structural vector autoregression of macroeconomic variables, in order to estimate the response of real out-

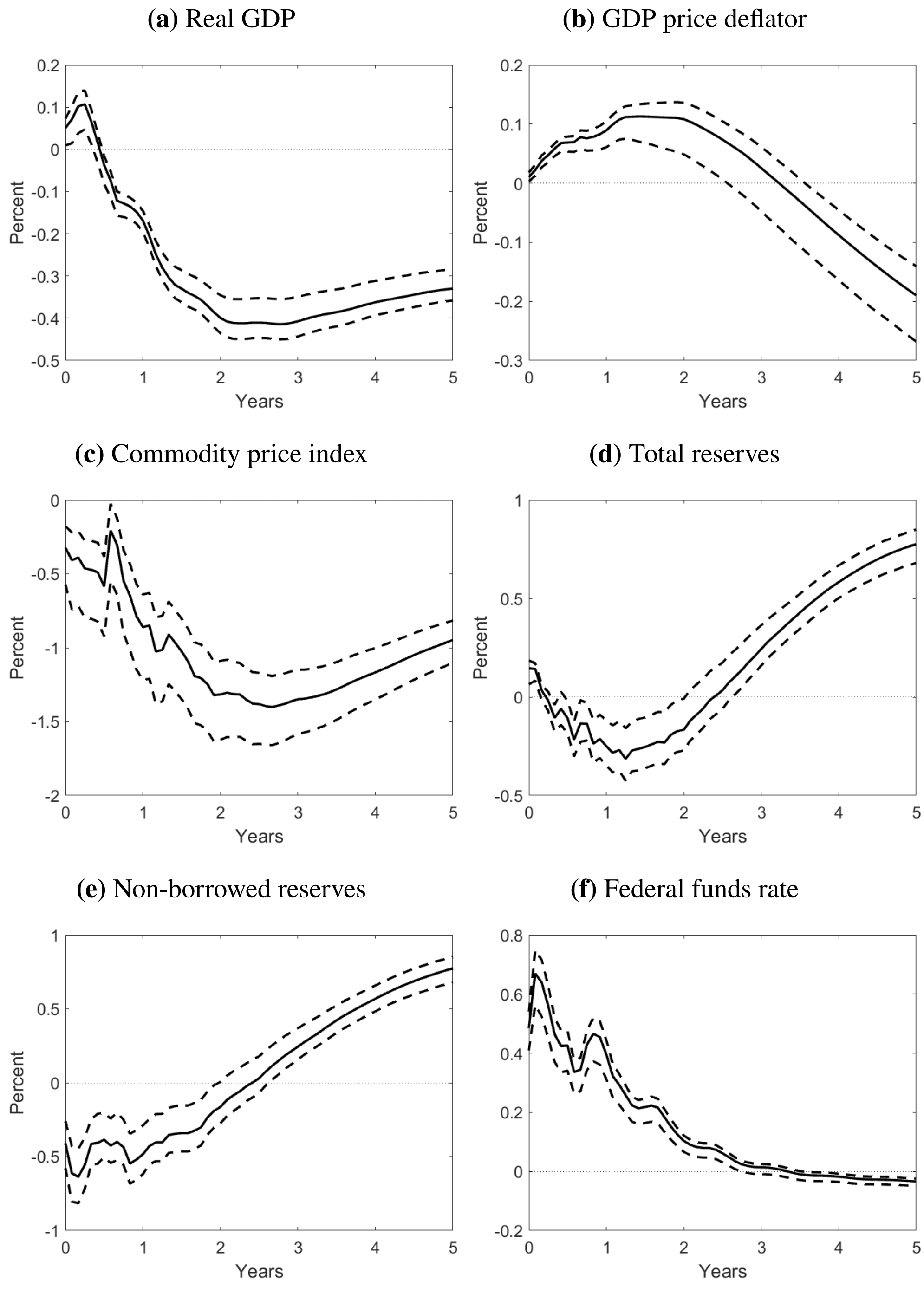

Figure 2: Responses to a contractionary monetary policy shock one standard deviation in size. The solid line represents the estimate for the real dataset. The upper and the lower dashed lines represent respectively the 84% quantile and the 16% quantile of the bootstrap estimates.

put and prices to a contractionary monetary policy shock. Identification is achieved by independent component analysis, since we assume that the structural shocks driving the macroeconomic dynamics are mutually independent.

In the spirit of Blanchard and Quah [1989], and appreciating the remarks by Lanne et al. [forthcoming], we have not imposed recursiveness as a short-run constraint on the causal relations among the observed variables. We have, instead, simplified the method in Hyvärinen et al. [2010] and Moneta et al. [2013], in order to use mutual independence as the only main assumption for identifying the structural vector autoregression. In our view, this new identification procedure narrows down economic *a priori* theorizing, and is loyal to the "agnostic" approach suggested by Uhlig [2005] and recently adopted by Arias et al. [2015].

Our empirical analysis has been conducted on the same United States macroeconomic dataset as in Uhlig [2005], covering the time span between January 1965 and December 2003. Results show that a negative influence of the contractionary monetary policy on real output materializes after around six months. Prices have a positive reaction for about three years, their response becoming negative thereafter.

References

J. E. Arias, D. Caldara, and J. F. Rubio-Ramirez. The systematic component of monetary policy in SVARs: An agnostic identification procedure. *International Finance Discussion Papers*, 2015(1131):1–37, mar 2015. doi: 10.17016/ifdp.2015.1131. URL `http://dx.doi.org/10.17016/ifdp.2015.1131`.

B. S. Bernanke and A. S. Blinder. The federal funds rate and the channels of monetary transmission. *American Economic Review*, 82(4):901–921, September 1992.

B. S. Bernanke and M. Gertler. Inside the black box: The credit channel of monetary policy transmission. *Journal of Economic Perspectives*, 9(4):27–48, nov 1995. doi: 10.1257/jep.9.4.27. URL `http://dx.doi.org/10.1257/jep.9.4.27`.

B. S. Bernanke and I. Mihov. Measuring monetary policy. *The Quarterly Journal of Economics*, 113(3):869–902, aug 1998. doi: 10.1162/003355398555775. URL `http://dx.doi.org/10.1162/003355398555775`.

D. A. Bessler and S. Lee. Money and prices: Us data 1869–1914 (a study with directed graphs). *Empirical Economics*, 27(3):427–446, 2002.

O. J. Blanchard and D. Quah. The dynamic effects of aggregate demand and supply disturbances. *The American Economic Review*, 79(4):655–673, 1989. ISSN 00028282. URL `http://www.jstor.org/stable/1827924`.

G. Chamberlain and M. Rothschild. Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica*, 51(5):1281, sep 1983. doi: 10.2307/1912275. URL `http://dx.doi.org/10.2307/1912275`.

L. J. Christiano, M. Eichenbaum, and C. Evans. The effects of monetary policy shocks: Evidence from the flow of funds. *The Review of Economics and Statistics*, 78(1):16, feb 1996. doi: 10.2307/2109845. URL `http://dx.doi.org/10.2307/2109845`.

L. J. Christiano, M. Eichenbaum, and C. L. Evans. Monetary policy shocks: What have we learned and to what end? In *Handbook of Macroeconomics*, pages 65–148. Elsevier BV, 1999. doi: 10.1016/s1574-0048(99)01005-8. URL `http://dx.doi.org/10.1016/s1574-0048(99)01005-8`.

P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3): 287–314, apr 1994. doi: 10.1016/0165-1684(94)90029-9. URL `http://dx.doi.org/10.1016/0165-1684(94)90029-9`.

C. Corradi-Dell'Acqua, A. Tusche, P. Vuilleumier, and T. Singer. Cross-modal representations of first-hand and vicarious pain, disgust and fairness in insular and cingulate cortex. *Nature Communications*, 7:10904, mar 2016. doi: 10.1038/ncomms10904. URL `http://dx.doi.org/10.1038/ncomms10904`.

S. Demiralp and K. D. Hoover. Searching for the causal structure of a Vector Autoregression. *Oxford Bulletin of Economics and Statistics*, 65(s1):745–767, dec 2003. doi: 10.1046/j.0305-9049.2003.00087.x. URL `http://dx.doi.org/10.1046/j.0305-9049.2003.00087.x`.

M. Eichenbaum. Comments on 'Interpreting the macroeconomic time series facts: The effects of monetary policy'. *European Economic Review*, 36(5):1001–1011, jun 1992. doi: 10.1016/0014-2921(92)90042-u. URL `http://dx.doi.org/10.1016/0014-2921(92)90042-U`.

A. M. Fraser and H. L. Swinney. Independent coordinates for strange attractors from mutual information. *Physical Review A*, 33(2):1134–1140, feb 1986. doi: 10.1103/physreva.33.1134. URL `http://dx.doi.org/10.1103/physreva.33.1134`.

A. Golan. Information and entropy econometrics - Editor's view. *Journal of Econometrics*, 107(1-2):1–15, mar 2002. doi: 10.1016/s0304-4076(01)00110-5. URL `http://dx.doi.org/10.1016/s0304-4076(01)00110-5`.

C. Gouriéroux, A. Monfort, and J.-P. Renne. Statistical inference for independent component analysis: Application to structural VAR models. *Journal of Econometrics*, forthcoming. doi: 10.1016/j.jeconom.2016.09.007. URL `http://dx.doi.org/10.1016/j.jeconom.2016.09.007`.

H. Herwartz and M. Plödt. The macroeconomic effects of oil price shocks: Evidence from a statistical identification approach. *Journal of International Money and Finance*, 61: 30–44, mar 2016. doi: 10.1016/j.jimonfin.2015.11.001. URL `http://dx.doi.org/10.1016/j.jimonfin.2015.11.001`.

J. Himberg and A. Hyvärinen. Icasso: software for investigating the reliability of ICA estimates by clustering and visualization. In *IEEE XIII Workshop on Neural Networks for Signal Processing (IEEE Cat. No.03TH8718)*. Institute of Electrical & Electronics Engineers (IEEE), 2003. doi: 10.1109/nnsp.2003.1318025. URL `http://dx.doi.org/10.1109/nnsp.2003.1318025`.

J. Himberg, A. Hyvärinen, and F. Esposito. Validating the independent components of neuroimaging time series via clustering and visualization. *NeuroImage*, 22(3):1214–

1222, jul 2004. doi: 10.1016/j.neuroimage.2004.03.027. URL `http://dx.doi.org/10.1016/j.neuroimage.2004.03.027`.

A. Hyvärinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634, may 1999. doi: 10.1109/72.761722. URL `http://dx.doi.org/10.1109/72.761722`.

A. Hyvärinen, K. Zhang, S. Shimizu, and P. O. Hoyer. Estimation of a Structural Vector Autoregression Model using non-Gaussianity. *Journal of Machine Learning Research*, 11:1709–1731, 2010.

S. Kullback and R. A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22(1):79–86, mar 1951. doi: 10.1214/aoms/1177729694. URL `http://dx.doi.org/10.1214/aoms/1177729694`.

M. Lanne and H. Lütkepohl. Structural vector autoregressions with nonnormal residuals. *Journal of Business & Economic Statistics*, 28(1):159–168, jan 2010. doi: 10.1198/jbes.2009.06003. URL `http://dx.doi.org/10.1198/jbes.2009.06003`.

M. Lanne, M. Meitz, and P. Saikkonen. Identification and estimation of non-Gaussian structural vector autoregressions. *Journal of Econometrics*, forthcoming. doi: 10.1016/j.jeconom.2016.06.002. URL `http://dx.doi.org/10.1016/j.jeconom.2016.06.002`.

A. Moneta. Graphical causal models and VARs: An empirical assessment of the real business cycles hypothesis. *Empirical Economics*, 35(2):275–300, 2008.

A. Moneta, D. Entner, P. O. Hoyer, and A. Coad. Causal inference by independent component analysis: Theory and applications. *Oxford Bulletin of Economics and Statistics*, 75(5):705–730, jul 2013. doi: 10.1111/j.1468-0084.2012.00710.x. URL `http://dx.doi.org/10.1111/j.1468-0084.2012.00710.x`.

C. J. Neely. The Federal Reserve responds to crises: September 11th was not the first. *Federal Reserve Bank of St. Louis Review*, 86(2):27–42, 2004.

S. Shimizu, P. O. Hoyer, A. Hyvärinen, and A. Kerminen. A linear non-Gaussian acyclic model for causal discovery. *The Journal of Machine Learning Research*, 7:2003–2030, 2006.

N. A. Siegfried. An information-theoretic extension to structural VAR modelling. Technical report, Hamburg University Quantitative Macroeconomics Working Paper 20203, 2002. URL `http://econwpa.repec.org/eps/em/papers/0203/0203005.pdf`.

C. A. Sims. Macroeconomics and reality. *Econometrica*, 48(1):1, jan 1980. doi: 10.2307/1912017. URL `http://dx.doi.org/10.2307/1912017`.

C. A. Sims. Are forecasting models usable for policy analysis? *Federal Reserve Bank of Minneapolis Quarterly Review*, pages 2–16, 1986.

C. A. Sims. Interpreting the macroeconomic time series facts. *European Economic Review*, 36(5):975–1000, jun 1992. doi: 10.1016/0014-2921(92)90041-t. URL `http://dx.doi.org/10.1016/0014-2921(92)90041-t`.

C. A. Sims and T. Zha. Were there regime switches in u.s. monetary policy? *American Economic Review*, 96(1):54–81, feb 2006. doi: 10.1257/000282806776157678. URL `http://dx.doi.org/10.1257/000282806776157678`.

J. H. Stock and M. W. Watson. Vector autoregressions. *Journal of Economic Perspectives*, 15(4):101–115, nov 2001. doi: 10.1257/jep.15.4.101. URL `http://dx.doi.org/10.1257/jep.15.4.101`.

H. Uhlig. What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52(2):381–419, mar 2005. doi: 10.1016/j.jmoneco.2004.05.007. URL `http://dx.doi.org/10.1016/j.jmoneco.2004.05.007`.

Appendix A. Different Data

Our main empirical analysis employs the same variables and the same time span as in Uhlig [2005] and in Arias et al. [2015]. In this appendix, instead, we consider the variables in Sims [1986]: real gross national product, real investment, GNP price deflator, the M1 measure of money, unemployment, and Treasury-bill rates (see Table A.1; "real inventories" will be added in an extended version of the dataset).[6]

As in Sims [1986], the variables are observed at quarterly frequency. We adopt the same time span of Uhlig [2005] and of our previous analysis, that is from year 1965 until year 2003, so that changes in results, between the main text and this appendix, cannot be attributed to the estimation window.

We consider the Treasury-bill rates as the monetary policy variable (taking the role played by the Federal funds rate in our previous analysis). Therefore, following the independent component analysis, we label as "monetary policy" shock the structural shock which affects Treasury-bill rates the most, and we call "contractionary" the corresponding one standard deviation impulse which increases the Treasury-bill rates.

The estimated impulse-response functions (i.e. the responses of each of the six variables to the contractionary monetary policy impulse) are represented in Figure A.1. The price puzzle emerges even more clearly than in our main analysis: the top-right panel of Figure A.1 indeed shows the GNP deflator responding positively immediately after the shock; the response becomes approximately flat after three years. The response of the real GNP (top-left panel of Figure A.1) becomes negative after about one year, a result that is not dissimilar from what we have previously obtained.

Interestingly, the initial response of the real GDP seems to be positive, as in our previous Figure 2 and in many previous studies (see e.g. the bottom-left panel of Figure 5 in Sims, 1992, p. 986, and, in symmetric representation, the top panels of Figure II in Bernanke and Mihov, 1998, p. 893). As pointed out by Bernanke and Gertler [1995, p. 32], inventory accumulation could partially explain the de-

[6]The data were retrieved from the FRED database of the Federal Reserve Bank of St. Louis, `https://fred.stlouisfed.org/` , on June 27, 2016. We have substituted the "business fixed investment" used by Sims [1986] with "real investment", as proxied by gross fixed capital formation.

Variable	File code in FRED
Real gross national product	GNPC96
GNP price deflator	GNPDEF
Real investment (gross fixed capital formation)	NAEXKP04USQ661S
M1 money stock	MANMM101USQ189S
Unemployment rate	USAURHARMQDSMEI
Interest rate T-bills	TB3MS
Real inventories (non-farm change in private inventories)	A015RX1Q020SBEA

Table A.1: Variables used in Appendix B, retrieved from the FRED database of the Federal Reserve Bank of St. Louis. All variables, except for real inventories, form our dataset D1. All variables including real inventories form our dataset D2.

layed reaction of real gross product to a contractionary monetary policy (aggregate production following demand only after some lag). Therefore, we now extend the dataset D1 by adding an inventory variable (see last row of Table A.1), and we rename the dataset as D2.

As can be seen in the bottom-left panel of Figure A.2, inventories are likely to increase following a contractionary monetary policy shock. After one year, the response becomes negative, and it approaches zero around three and a half years following the shocks (that is when the GNP response becomes straighter, see top-left panel of Figure A.2). The inclusion of inventories in the analysis does not alter our main findings: the real output responds negatively to the contractionary shock in the medium-long run, while prices increase at least in the short-medium run.

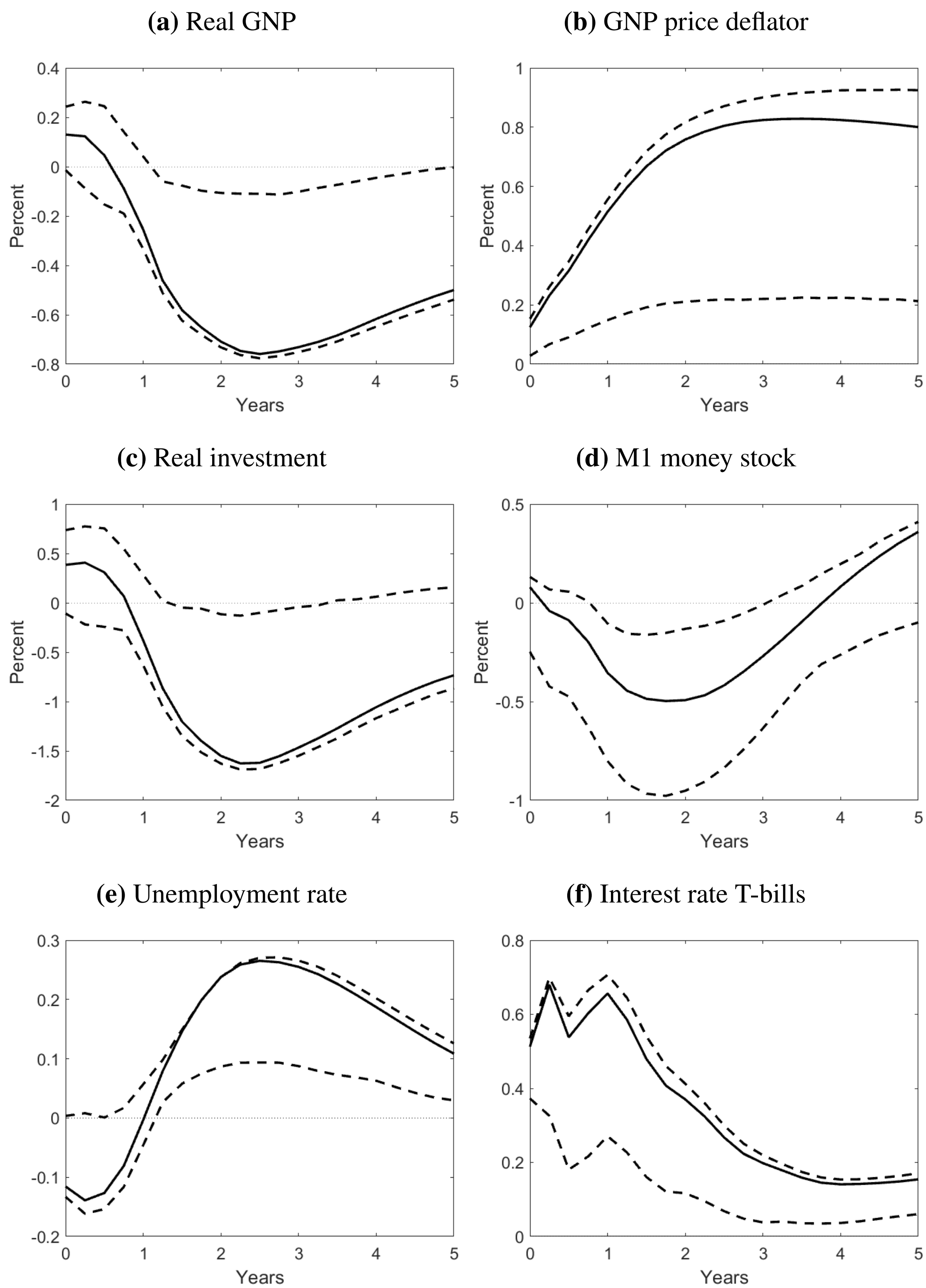

Figure A.1: Responses to a contractionary monetary policy shock one standard deviation in size, estimated using the dataset D1 (quarterly data; same variables as in Sims, 1986). The solid line represents the estimate for the real dataset. The upper and the lower dashed lines represent respectively the 84% quantile and the 16% quantile of the bootstrap estimates.

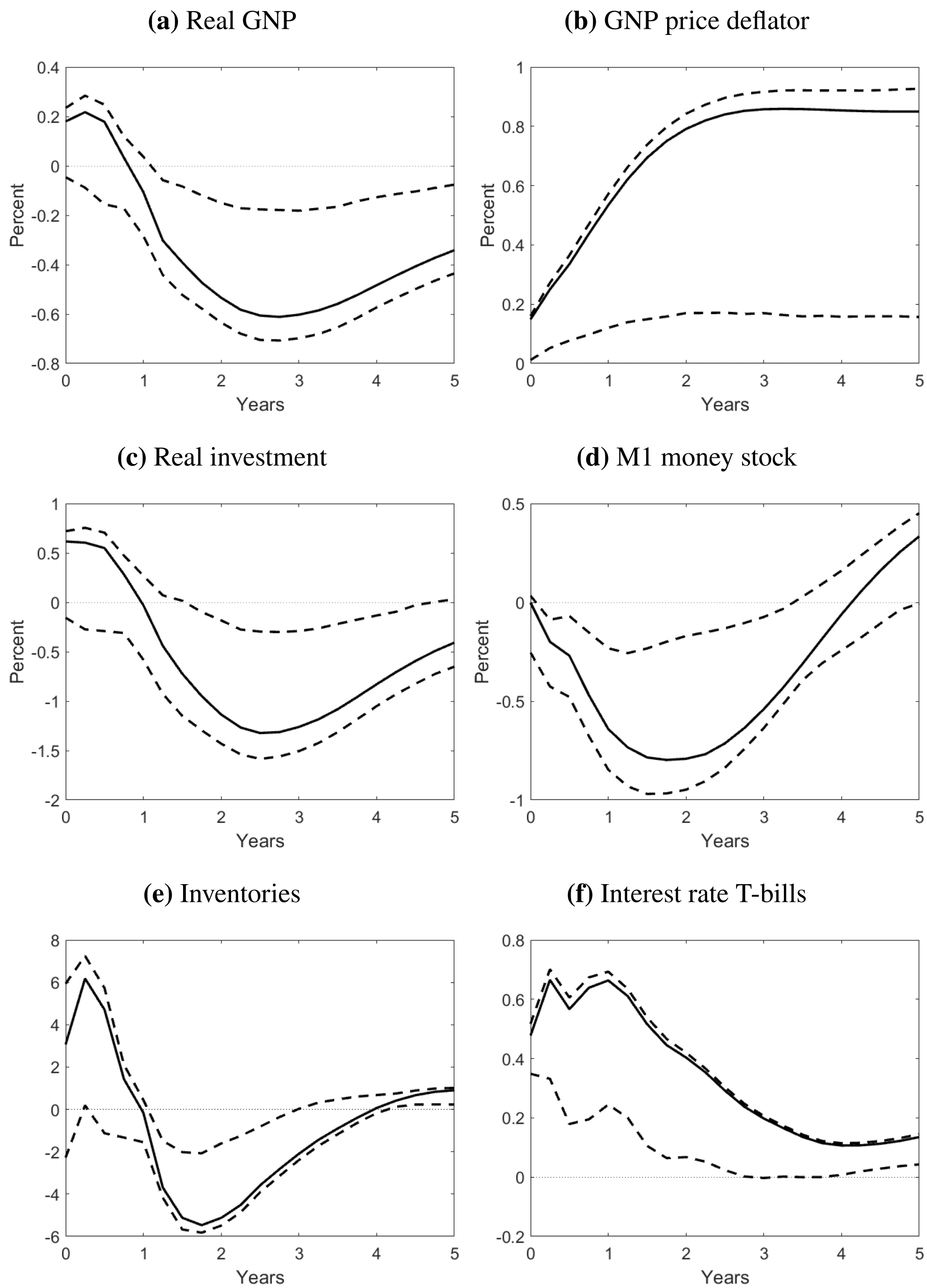

Figure A.2: Responses to a contractionary monetary policy shock one standard deviation in size, estimated using the dataset D2 (quarterly data; same variables as in Sims, 1986 plus inventories). The solid line represents the estimate for the real dataset. The upper and the lower dashed lines represent respectively the 84% quantile and the 16% quantile of the bootstrap estimates. The response of the unemployment rate is not shown.

Appendix B. Different Shock Labels

In our empirical analysis, the structural shock affecting the Federal funds rate the most, at impact, has been labelled as "monetary policy" shock. In this appendix, we analyze the response of the observable variables to two other structural shocks, exploring the possibility that monetary policy could be associated to a different shock, or to more than one shock. Among the observable variables, we consider total reserves and non-borrowed reserves as the two other possible policy instruments (apart from the Federal funds rate) which are under a direct influence of the central bank.

First, we define the "labelling scheme" A1 according to the hypothesis that a "monetary policy" shock affects total reserves the most. In the mixing matrix previously reported in Table 1, the first column contains the highest number, in absolute value, of the fourth row (the row associated to total reserves). Therefore, under the labelling scheme A1, the shock ε_1 would be labelled as a "monetary policy" structural shock.
The responses of the observed variables to shock ε_1 (one standard deviation impulse, lowering total reserves at impact) are shown in Figure B.1. In particular, the response of GDP (top-left panel) is positive at all time lags, the response of the GDP deflator (top-right panel) is negative for the first three years, and the response of the commodity price index (center-left panel) is positive after four years. These results would let us think that the labelling scheme A1 does not properly recognize the monetary policy shock. Rather, it might wrongfully label, as monetary policy shock, a systematic response of monetary policy to output booms and recessions. Moreover, the confidence intervals in Figure B.1 indicate a high variance and asymmetry of the response estimates. This could be due to the presence, in the datasets, of abrupt changes in total reserves, unexplainable through the usual connections of the total reserves to the rest of the economic system (an exemplary event is September 2001: see pages 37-39 of Neely, 2004, and page 62 of Sims and Zha, 2006).

The labelling scheme A2 assumes, instead, that non-borrowed reserves are the main monetary policy instrument. However, non-borrowed reserves may decrease as a consequence of the decrease in total reserves, which we have considered as main policy instrument under the labelling scheme A1. Therefore, we now point at the structural shock which affects non-borrowed reserves the most, at impact, only after excluding the shock affecting total reserves the most. Considering the mixing matrix in Table 1, the structural shock ε_2 is thus now labelled as a "monetary policy" shock, since, after excluding the first column (associated to ε_1, with high impact on total reserves), the second column contains the highest number, in absolute value, of the fifth row (the row associated to non-borrowed reserves).

The responses of the observed variables to the shock ε_2 (one standard deviation impulse, lowering non-borrowed reserves at impact) are shown in Figure B.2. Within two-three months following the contractionary shock, the real output declines; it begins to recover only after one-two years (possibly due to a fast decrease of the Federal funds rate). The commodity price index drops immediately, and the response remains negative almost uninterruptedly for at least two years. The response of the GDP price deflator, after an uncertain period of about one year, becomes markedly and constantly negative. Putting together the results illustrated in the different panels Figure B.2, we can conclude that the labelling scheme A2 provides a credible alternative to the benchmark scheme (where the Federal funds rate is the main policy target). Indeed, the scheme A2 highlights an alternative monetary channel exploited by policy makers, with non-borrowed reserves used as the main policy instrument, and interest rates playing only a minor role (at least in the short run). Interestingly, no price puzzle would emerge from the responses to the structural shock labelled as "monetary policy" under this alternative scheme.

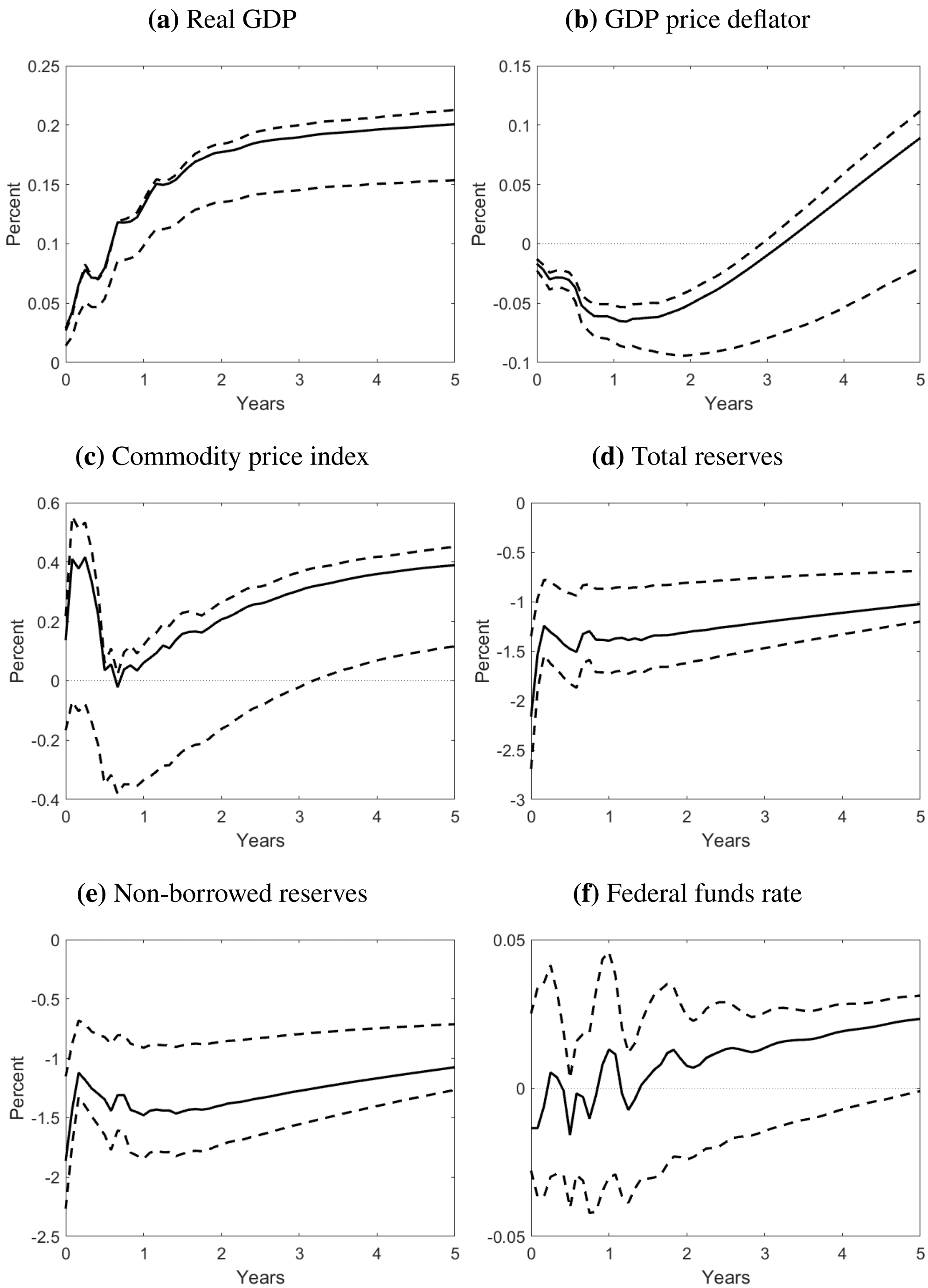

Figure B.1: Responses to a contractionary monetary policy shock one standard deviation in size, under the labelling scheme A1 (monetary policy shock impacting mainly total reserves). The solid line represents the estimate for the real dataset. The upper and the lower dashed lines represent respectively the 84% quantile and the 16% quantile of the bootstrap estimates.

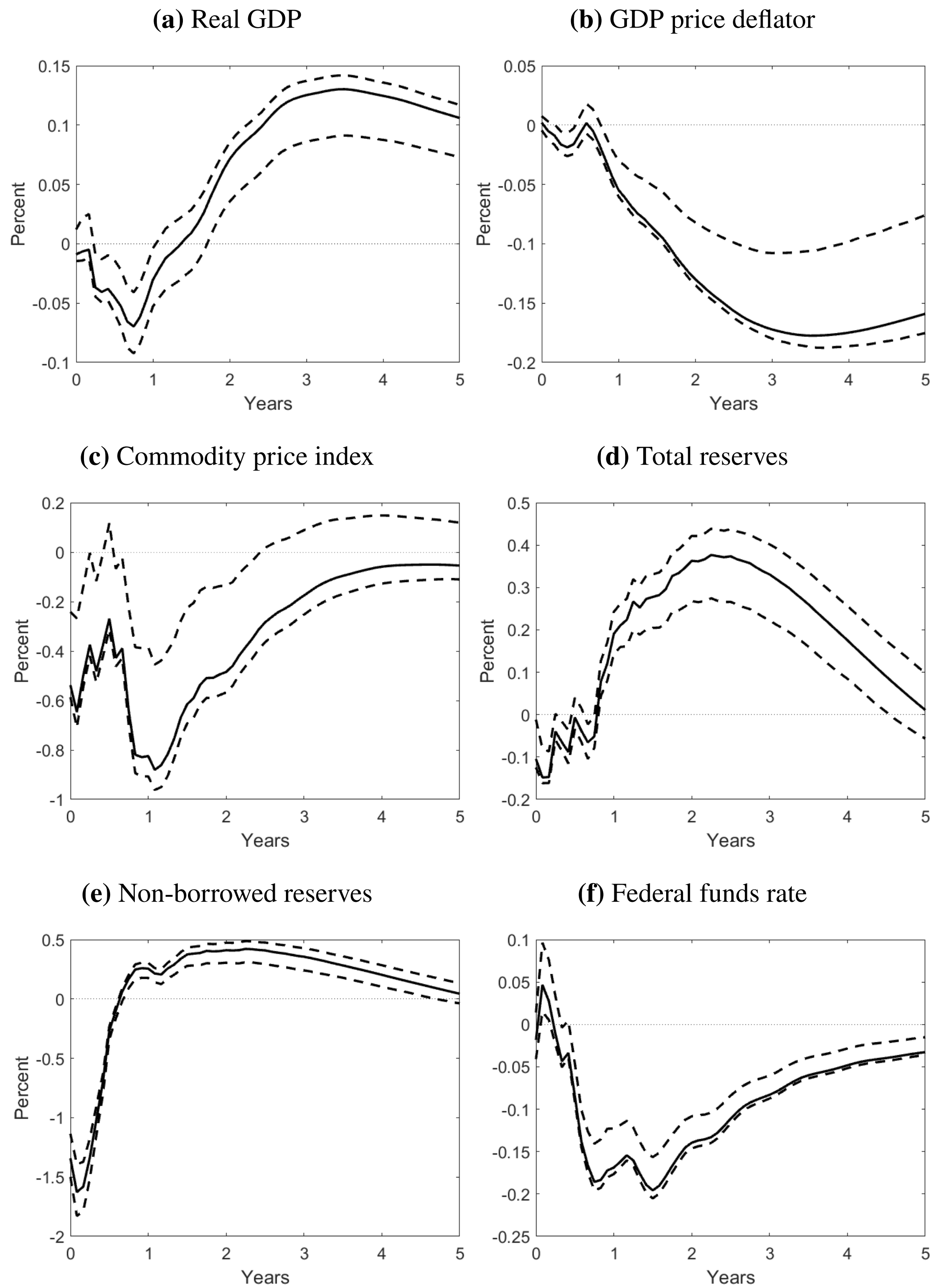

Figure B.2: Responses to a contractionary monetary policy shock one standard deviation in size, under the labelling scheme A2 (monetary policy shock impacting mainly non-borrowed reserves). The solid line represents the estimate for the real dataset. The upper and the lower dashed lines represent respectively the 84% quantile and the 16% quantile of the bootstrap estimates.