NEW WAYS TO ESTIMATE THE DYNAMIC EFFECTS OF MACROECONOMIC POLICIES

by

Balázs Vonnák

Submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy at
Central European University,
Department of Economics and Business

Supervisor: Róbert Lieli

Budapest, Hungary

10.14754/CEU.2021.04

CENTRAL EUROPEAN UNIVERSITY
DEPARTMENT OF ECONOMICS AND BUSINESS

The undersigned hereby certify that they have read and recommend to the Department of Economics and Business for acceptance a thesis entitled **"New Ways to Estimate the Dynamic Effects of Macroeconomic Shocks"** by Balázs Vonnák.
Dated: September 20, 2021

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Chair of the Thesis Committee:

DocuSigned by:
Mátyás László
F608CAA1BA114DB...
László Mátyás

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Advisor:

DocuSigned by:
Robert Lieli
243BCFE3F32841C...
Robert Lieli

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Internal Examiner:

DocuSigned by:
Sergey Lychagin
0B1F487DF976478...
Sergey Lychagin

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
External Examiner:

DocuSigned by:
Giovanni Ricco
1566A6D4494F4EC...
Giovanni Ricco

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
External Member:

DocuSigned by:
Kőrösi Gábor
90C598720A304CC...
Gábor Kőrösi

10.14754/CEU.2021.04

Central European University
Department of Economics and Business

Author: Balázs Vonnák
Title: New Ways to Estimate Dynamic Effects of Macroeconomic Policies
Degree: Ph.D.
Dated: October 4, 2021

Hereby I testify that this thesis contains no material accepted for any other degree in any other institution and that it contains no material previously written and/or published by another person except where appropriate acknowledgement made.

signature of the author

Abstract

In the first chapter I show how the generalized propensity score can be used to estimate dynamic causal effects. This approach is flexible in the sense that the effect can depend on the level of the policy variable and its change in any nonlinear way. Compared to the local projection, the propensity score based estimator has much fewer parameters to be estimated, which is a favourable property when the times series are short. A very general Monte Carlo simulation reveals that it yields more precise impulse responses than the VAR and the local projection when potential misspecification and low degrees of freedom is an important issue.

In the second chapter I apply the propensity score based estimator to U.S. data to investigate how the effect of the Fed's interest rate changes depends on the initial level of the interest rate and on the size of the change. Although exogenous shocks to the interest rate are estimated with a linear SVAR model, the results display significant nonlinearities. Most importantly, monetary policy decisions seem to exert faster and larger impact on GDP and consumer prices when the interest rate is high. On the other hand, dependence on the size and the direction of the interest rate change does not show signs of nonlinearity.

In the third chapter a new instrument for monetary policy shocks is proposed. Exogenous variation of the policy rate may come from frictions of collective decision-making. Dissenting votes indicate how far the final decision of the decision making body is from the mean of the members' individually preferred interest rates and thus correlate with the policy shocks caused by the decision-making frictions. Measures of dissent are used as external instrument in a structural VAR. Results for the U.S. show significant effect

10.14754/CEU.2021.04

of the Fed's interest rate policy on real variables with the expected sign. On the other hand, the estimated effect on nominal variables is reminiscent of the price puzzle. Usual remedies, such as inclusion of commodity prices, inflation expectations or starting the sample in the middle of the eighties do not change the qualitative results casting doubt on the usual interpretation that the price puzzle is a statistical artifact.

Acknowledgements

I am deeply indebted to my supervisor Róbert Lieli for his continuous guidance on how to conduct my research and how to complete the thesis.

I also want to thank my examiners, Giovanni Ricco and Sergey Lychagin for their very useful comments, suggestions that helped to improve my thesis a lot.

I have benefited a lot from the discussions at and after the Brown Bag Seminars at the Central European University, where I received valuable feedback from the members of the Department of Economics and Business and from other PhD candidates.

I am also grateful to the participants of the 2021 International Association for Applied Econometrics Annual Conference, where I had the opportunity to present the first two chapters of my thesis, for their inspiring comments.

Contents

List of Figures

10.14754/CEU.2021.04

List of Tables

Chapter 1

Estimation of Dynamic Causal Effects with the Generalized Propensity Score

1.1 Introduction

Estimating the dynamic effects of macroeconomic shocks has been one of the most intensively explored area in time series econometrics during the past several decades. The most widely used tool for this purpose is the vector autoregressive model (VAR), advocated in the seminal paper of Sims (1980). VARs model the joint dynamics of the endogenous variables in a linear way. In order to identify causal effects, one has to formulate restrictions that may come from a formal theoretical model, or, as in most cases, from informal economic intuition. An appealing feature of identified VARs (or structural VARs, SVARs) is that the same model is used for identifying structural shocks as well as for calculating their dynamic effect on the endogenous variables (impulse response functions).

One of the drawbacks of SVARs is that they predict outcomes only one period ahead directly. In order to forecast for longer horizons, either unconditionally or conditionally on a structural shock, one has to iterate one-step-ahead predictions. During the iterations, prediction errors due to possible misspecification may be amplified resulting in imprecise estimates at longer horizons. Another limitation is that VARs are linear, therefore not capable of identifying nonlinear effects.

Alternatively, Jordà (2005) proposes direct regression to estimate the impulse response function. The outcome variables at period $t+h$ are regressed on variables at period $t, t-1, ...$, thus no iteration is needed for making projections. Local projections made this way are flexible in the sense that by including powers and interactions of the explanatory variables nonlinear impulse responses can be estimated easily, which is not possible with the linear VAR. Jordà (2005) also demonstrates that local projections are more robust to misspecification than VARs.

In this chapter I propose another way to estimate possibly nonlinear effects. The estimator is based on the concept of the propensity score, and makes use of its dimension reducing property. Thus it offers a powerful alternative to local projection if the number of observations is low, a typical problem when working with macroeconomic time series.

There are other approaches for estimating nonlinear impulse responses. Time varying VARs, such as that of Cogley and Sargent (2003) or Primiceri (2005) allow the parameters to change over time, thus, the impulse responses may differ across periods. However, for a certain period, the impulse responses are linear in the sense that they depend on the size (and sign) of the shock in a linear way. The nonlinearity is thus exogenously given, there is no explicit dependence on observable variables.

State dependence can be captured with smooth transition models. The smooth transition VAR of Weise (1999) or Rahman and Serletis (2010) allows the coefficients to depend on the level of endogenous variables. Using a smooth transition local projection model, Tenreyro and Thwaites (2016) estimate the asymmetric effect of monetary policy. The limitation of these models is that they suffer from the curse of dimensionality when applied to short time series. Another limitation is that these models can handle only few states, typically two. They are not appropriate for modelling state dependence with large number of possible states.

A general problem with VARs and local projections is that one has to estimate a huge number of parameters, which is an effective constraint when working with short time series or trying to identify nonlinearities. There are some very recent developments in coping with the problem of low degrees of freedom. Barnichon and Matthes (2018) propose a basis function approximation for estimating the moving average representation of the VAR. Basis function approximation reduces the dimensionality, therefore, may be useful for estimating nonlinearities even from short time series.

Angrist et al (2018) showed first how the propensity score can be adopted in a time series framework. They investigate the effect of the Fed's interest rate decisions on key macroeconomic variables by estimating a policy reaction function at the first stage, and then using the inverse probability weighting estimator for the outcomes. An important limitation of their identification strategy is that it is specific to the Fed's regular rate setting schedule and exploits the discrete nature of policy rate changes. Their method cannot be adopted when the policy variable can take many values.

My proposed approach is a generalization of the idea of Angrist et al (2018) to the continuous case. First, I show how the generalized propensity score (GPS) introduced in Hirano and Imbens (2004) can be used when the policy is described by a continuous variable. It has the same advantage as the propensity score typically used for binary treatment settings, namely that it can reduce the dimensionality substantially. The estimator based on the GPS can thus be useful when working with small sample, especially if we are interested in nonlinear effects.

Then I investigate the sampling properties of the proposed estimator. This is an exercise missing from Angrist et al (2018). I compare the performance of the GPS-based estimator to that of the most commonly used methods, namely, the VAR and the local projection. Since I focus on estimating the dynamic effect of a shock, the shocks are identified by the SVAR first, and then used by the other two approaches, too. In the Monte Carlo simulation I use a huge number of data generating processes (DGPs), all of them estimated from real data, in order to get a robust picture about the relative preciseness of the three models.

I find that the GPS-based estimator performs significantly better than Jordà's local projection and in most of the cases better than the VAR. Its advantage is more obvious when the time series are short, which is typical in empirical studies using macro data. When the DGP is linear and the VAR is correctly specified, the GPS-based estimator's prediction errors are comparable with that of the VAR in large samples, and within a two year horizon, in small samples, too. When the VAR is misspecified, the GPS-based estimator outperforms not only the VAR, but also the local projection. Finally, when the DGP is a regime switching VAR, the GPS-based estimator is the best estimator for a two-year horizon. Based on the simulation results, the estimator proposed in this chapter seems to offer a better alternative to local projection, especially when the number of observations is low.

The chapter is organized as follows. First, I sketch out the idea how the generalized propensity score can be used for estimating impulse responses from time series data. Then I show how the estimation is implemented. In section 1.4 the Monte Carlo simulation results are presented. Section 1.5 concludes.

1.2 Theoretical background

In this section I first present the basic idea of the generalized propensity score based estimator. Then I show how it relates to other widely used approaches such as the structural vector autoregression and the local projection with the help of a particular simple data generating process.

1.2.1 The basic concept

Let I_t denote the value of the policy variable at period t, Y_{t+h} $(h = 1, 2, ..., H)$ the scalar outcome h period later we are interested in, and $X_t = \{Z_{t-1}, Z_{t-2}, ..., e_t\}$ the set of all relevant covariates, where Z_t is the vector of observable variables and e_ts are the unobserved non-policy shocks. Note that the vector Z_t includes the policy (I_t) and the outcome variable (Y_t), too. We would like to estimate the effect of an exogenous change in the policy variable on the outcome variable $1, 2, ..., H$ periods later, that is the impulse response function.

VARs capture the one-step-ahead dependence of endogenous variables on their past values[1]. Let $f_{VAR}(Z_t, Z_{t-1}, ...) = E(Y_{t+1}|Z_t, Z_{t-1}, ...)$ denote this (linear) relationship. The underlying model behind the calculation of the impulse response function is the following:

$$E(Y_{t+h}|Z_t, Z_{t-1}, ...) = \underbrace{f^Y_{VAR}(f_{VAR}...(f_{VAR}}_{h\,\text{times}}(Z_t, Z_{t-1}, ...))...)$$

where f^Y_{VAR} denotes the element of the vector f_{VAR} that corresponds to the outcome variable Y. Since $f_{VAR}(Z_t, Z_{t-1}, ...)$ is a reduced form, it can be estimated by OLS. Projections for any horizons can be made by iterating one-step-ahead forecasts.

With local projection introduced in Jordà (2005), calculating impulse response functions for h periods ahead requires h separate estimations. For any horizon h, the relationship between the outcome variable and the explanatory variables $f_{LP,h}(Z_t, Z_{t-1}, ...)$ can be estimated by OLS. The impulse response function is constructed from the predictions of those regressions based on the following model:

$$E(Y_{t+h}|Z_t, Z_{t-1}, ...) = f_{LP,h}(Z_t, Z_{t-1}, ...)$$

The estimator proposed in this chapter is based on Hirano and Imbens (2004), who introduced the concept of the generalized propensity score (GPS), which is the analogue of the propensity score for continuous treatments, and is used for estimating so-called dose response functions[2]. In general, the GPS is the conditional density of a continuous treatment variable given a set of covariates, denoted by $r(i, x) = f_{I|X}(i|x)$. In the present context I define the GPS as the conditional density of the treatment (policy) variable I_t given the vector of covariates X_t or, loosely speaking, the probability of a particular policy given the state of the economy at period t. Note that this conditional density depends on time only through the covariates. The GPS associated with the level of the

[1]And possibly on exogenous variables, but for the sake of simplicity, we ignore them here.
[2]A similar approach is presented in Imai and van Dyk (2004)

policy variable in period t is then

$$R_t = r(I_t, X_t).$$

The key condition for removing the endogeneity bias is weak unconfoundedness, which is formulated as follows with $Y_{t+h}(i)$ denoting the potential outcome after a policy i:

$$Y_{t+h}(i) \perp I_t | X_t \text{ for all } i \tag{1.1}$$

Weak unconfoundedness means that once we control for all the relevant confounders, potential outcomes h periods after a given policy are orthogonal to the actual policy observed in period t, and this is true for all hypothetical policies i. The qualifier "weak" stresses that we do not require the joint distribution of potential outcomes after all potential policies to be independent of the actual policy.

Intuitively, weak unconfoundedness means that if we compare two periods in which the economy in exactly the same state, the actual policy is independent of the future potential outcomes, that is, differences in the observed policies can be treated as being random. In practice the assumption essentially is that all relevant confounders are taken into account during estimation.

As Hirano and Imbens (2004) show, weak unconfoundedness implies that it is enough to control for the generalized propensity score to ensure that the probability distribution of possible policies is (conditionally) independent of the potential outcomes related to that policy, or formally

$$f_I(i|r(i, X_t), Y_{t+h}(i)) = f_I(i|r(i, X_t))$$

for every i.

It follows from this result that instead of controlling for all the covariates that influence both the policy response and the future path of the outcome variable, it is sufficient to control only for the generalized propensity score in order to eliminate the endogeneity bias, that is

$$Y_{t+h}(i) \perp I_t | R_t \text{ for all } i.$$

As a consequence, to estimate the effect of a policy change with the generalized propensity score, the outcomes (Y_{t+h}) have to be regressed only on the policy variable (I) and the estimated propensity score associated with that policy (R). Obtaining the GPS requires a first stage estimation, to be described later. The dependence of the outcome on the policy variable and the GPS ($f_{GPS,h}$) can be estimated by OLS, and similarly to local projections, requires h separate estimations for each horizons of interest. The projection for horizon h is then calculated directly from those regression estimates based on the model:

$$E(Y_{t+h}|Z_t, Z_{t-1}, ...) = f_{GPS,h}(I_t, R_t)$$

The obvious advantage here is that one has to control only for two variables at t, rather than for all confounders at $t, t-1, ...$ as in the other two cases. Any nonlinear approximation can be made easily by including higher order terms without decreasing the degrees of freedom substantially.

The cost of this reduction of dimensionality is the need of a first stage estimation of the GPS which necessarily adds to the total estimation error. Moreover, estimating the GPS itself may require a high-dimensional model to identify the policy shocks, thus, seemingly we cannot get rid of the curse of dimensionality, but generally this is not true. As long as the policy rule can be well approximated by a linear model, one can obtain reasonably precise estimates of the policy shocks and, thus, the propensity score. For example, a linear VAR fitted on a non-linear vector process can yield good estimates of the one-step ahead forecast errors that are eventually used for estimating the structural shocks. Indeed, this is confirmed by the simulation results to be shown later.

10.14754/CEU.2021.04

1.2.2 A simple model

To see how the GPS-based estimation relates to SVAR and local projection, consider the following data generating process:

$$Y_t = \alpha I_{t-1} + e_t$$

$$I_t = \beta Y_{t-1} + \varepsilon_t$$

where I is the policy variable, Y is the outcome, ε and e are exogenous, orthogonal, i.i.d. shocks to the two variables. Note that this is a very special structural VAR model, in which the two structural shocks enter the two equations separately, and ε is the exogenous variation in the policy variable.

The impulse response of Y to one unit shock to I at period 0 is shown in Table 1.1.

time	0	1	2	3	4	5	6
Y	0	α	0	$\alpha^2\beta$	0	$\alpha^3\beta^2$	0

Table 1.1: Impulse response of Y to one unit shock to I at period 0

Within the VAR framework, one estimates the following model with OLS:

$$Z_t = AZ_{t-1} + v_t,$$

where $Z_t = (Y_t, I_t)'$, $v_t = (e_t, \varepsilon_t)$. Since the VAR is correctly specified, the estimator is consistent, and thus asymptotically

$$\hat{A} = \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}.$$

The VAR estimate of the impulse response function as well as that of the policy shocks will be consistent, too. The first one is the upper right element of $\hat{A}^h$, the latter is $\hat{\varepsilon}_t = I_t - \hat{\beta}Y_{t-1}$.

With local projection, the regression model for the response h period ahead is the following:

$$Y_{t+h} = a_y^h Y_t + a_i^h I_t + w_t^h.$$

By iteration, the true relationships are

$$Y_{t+1} = \alpha I_t + e_{t+1}$$

$$Y_{t+2} = \alpha\beta Y_t + \alpha\varepsilon_{t+1} + e_{t+2}$$

$$Y_{t+3} = \alpha^2\beta I_t + \alpha\beta e_{t+1} + \alpha\varepsilon_{t+2} + e_{t+3}$$

$$...,$$

and since the linear combinations of future realizations of the structural shocks are independent of Y_t and I_t, the regression models used for local projection are also correctly specified[3], and thus the impulse response function, which is the sequence $\{\hat{a}_i^1, \hat{a}_i^2, \hat{a}_i^3, ...\}$ is consistently estimated.

To see how GPS-based estimation is carried out, let $r(x|Z_{t-1})$ be the conditional density function of I_t and $g(x)$ the density function of ε_t. Note that in this very simple example the mean of I_t depends only on Y_{t-1}, but generally all the other endogenous variables and more lags should be included, too. Then the generalized propensity score of observation t is

$$R_t = r(I_t|Z_{t-1}) = g(I_t - \beta Y_{t-1}) = g(\varepsilon_t).$$

Consistent estimates of ε_t and β can be obtained, for example, from the VAR. Therefore, the GPS can be derived for each observation as $\hat{R}_t = g(I_t - \hat{\beta}Y_{t-1})$. To calculate the impulse response function for horizon h, one has to regress Y_{t+h} on $I_t, g(I_t - \hat{\beta}Y_{t-1}), I_t^2, I_t g(I_t - \hat{\beta}Y_{t-1}), g(I_t - \hat{\beta}Y_{t-1})^2, ...$ At this stage an approximation of $f_{GPS,h}$ is needed.[4]

It is also worth seeing what the weak unconfoundedness assumption's implications are. Let us focus on the one period ahead forecast, that is on

$$Y_{t+1} = \alpha I_t + e_{t+1}.$$

[3]Of course, the residual term w_t^h may now be autocorrelated.

[4]More details about the implementation are given in the next section.

For any hypothetical interest rate i at period t, the conditional covariance between the potential outcomes one period later ($Y_{t+1}(i)$) and the actual interest rate at t is

$$\mathrm{Cov}(Y_{t+1}(i), I_t|Y_{t-1}) = \mathrm{Cov}(\alpha i + e_{t+1}, \beta Y_{t-1} + \varepsilon_t|Y_{t-1}) =$$

$$= \mathrm{Cov}(e_{t+1}, \varepsilon_t|Y_{t-1}) = \mathrm{Cov}(e_{t+1}, \varepsilon_t),$$

which is zero because of the orthogonality assumptions for the structural shocks, that is, weak unconfoundedness is true for one period ahead. The same can also be shown for longer horizons similarly.

Obviously, in this example the proposed estimator has no expected advantage over VAR or local projection, because they capture the data generating process correctly. Here I only wanted to demonstrate through a simple example how estimating an impulse response with the generalized propensity score relates to the benchmark approaches. Interestingly, as shown in Section 1.4, the estimation with the GPS performs almost as well as the correctly specified linear models even in this case.

Generally, the estimation of the policy shocks is not as straightforward as in this example and involves the imposition of identifying restrictions. But once the necessary restriction are imposed, the SVAR estimates of the coefficient matrix and the impact response matrix can be used to calculate the historical policy shocks that are a linear function of the observed variables. This linear relationship is essentially determined by the systematic behaviour of the policy-maker. The main advantage of the GPS-based estimator over local projection is that lagged endogenous variables appear in the final regressions only in a fixed linear combination through the density function (in the previous example as $g(I_t - \hat{\beta}Y_{t-1})$) that determines the propensity score. This property helps save degrees of freedom when estimating the dynamic effect of policy changes on future outcomes.

1.3 Estimation strategy

First I show how dynamic causal effects can be generally identified by using the generalized propensity score. Then I describe the main steps of the estimation procedure.

1.3.1 Identification

Let the law of motion be

$$Z_t = \Gamma_1(Z_{t-1}, Z_{t-2}, ..., e_t, \varepsilon_t). \tag{1.2}$$

where ε_t is the policy shock, or the exogenous variation of the treatment, and e_t is the vector of the other structural shocks. (1.2) can be regarded as a generalized form of a (S)VAR, that is the endogenous variables depend on their past values and contemporaneous exogenous shocks, but Γ_1 is not necessarily linear. The shocks are assumed to be serially and mutually independent.

If the policy rule[5] is

$$I_t = \Phi(Z_{t-1}, Z_{t-2}, ..., e_t, \varepsilon_t), \tag{1.3}$$

the weak unconfoundedness assumption (1.1) can be written as

$$Y_{t+h}(i) \perp \varepsilon_t | Z_{t-1}, Z_{t-2}, ..., e_t \text{ for all } i \text{ and } h = 1, 2, ..., H \tag{1.4}$$

which means that once we control for the past state of the economy and current shocks not related to the policy, the actual realization of the current policy shock is independent of what the effects of a particular policy can be. Note that $\{Z_{t-1}, Z_{t-2}, ..., e_t\}$ is the collection of all potential confounders, denoted earlier by X_t.

Similarly to Jordà's (2005) local projection, the estimator I propose in this chapter is not capable of identifying the policy shocks in itself, only their dynamic effect. There are several ways to identify exogenous variation of the policy variable. If additive separability

[5]Note that (1.3) is one equation of the system (1.2).

of the policy shock in the reaction function (1.3) is a reasonable assumption, that is

$$I_t = \Phi'(Z_{t-1}, Z_{t-2}, ..., e_t) + \varepsilon_t,$$

one can estimate Φ', and derive the policy shocks as regression residuals. The problem is that in most cases this is infeasible because e_t-s are unobserved.

The other way is to use identifying restrictions within a VAR framework, that is a structural VAR. In particular, if Φ is linear, (1.3) is equivalent to the structural decomposition of the corresponding equation of a VAR. The advantage of using SVARs for identification is that there is a huge literature on how to estimate idiosyncratic shocks of monetary, fiscal and other policies within that framework. Since in the simulation exercise I will estimate the policy shocks by a SVAR, it is useful to see how its assumptions relate to those of the propensity score based identification.

It follows from (1.4) that to achieve unconfoundedness, the VAR should contain all the relevant confounders, and the identifying restrictions should separate the (unobserved) policy shocks from the (unobserved) other shocks. To see how unconfoundedness can be provided, let us write the vector of the observable variables at $t + h$ as a function of observable variables up to t and structural shocks thereafter, using (1.2):

$$Z_{t+h} = \Gamma_h(Z_t, Z_{t-1}, ..., e_{t+1}, ..., e_{t+h}, \varepsilon_{t+1}, ..., \varepsilon_{t+h})$$

The potential outcomes associated with policy i are then

$$Z_{t+h}(i) = \Gamma_h(\{i, X_t(i)\}, Z_{t-1}, Z_{t-2}, ..., e_{t+1}, ..., e_{t+h}, \varepsilon_{t+1}, ..., \varepsilon_{t+h}),$$

and since $X_t(i)$ is a function of $Z_{t-1}, Z_{t-2}, ..., e_t, i$, which follows from (1.3) and (1.2), we have

$$Z_{t+h}(i) = \Gamma'_h(i, Z_{t-1}, Z_{t-2}, ..., e_t, e_{t+1}, ..., e_{t+h}, \varepsilon_{t+1}, ..., \varepsilon_{t+h}),$$

Thus, for any policy i, once we keep all confounders $Z_{t-1}, Z_{t-2}, ..., e_t$ fixed, the potential outcomes are a function of future shocks $e_{t+1}, ..., e_{t+h}, \varepsilon_{t+1}, ..., \varepsilon_{t+h}$, while the actual policy is a function of only ε_t, according to (1.3). Weak unconfoundedness is now equivalent to the independence of current policy shocks from all future structural shocks which is typically provided by definition, in the SVAR framework, too.

Of course, to meet the weak unconfoundedness assumption, large number of controls might be needed when identifying exogenous variation of the policy variable. But, as mentioned earlier, if at this stage a linear approximation is good enough to characterize the reaction function, the GPS-based approach requires much fewer parameters to be estimated.

It is worth noting that this approach can be used to estimate the effect of any macroeconomic shock such as aggregate demand, technology etc., not only that of a policy shock, even if there is no obvious "treatment" variable, there is no policy-maker. Let us consider the example of aggregate demand. Shocks to aggregate demand move the price level and output in the same direction, according to conventional macro theories. Of course, prices and output are determined not only by demand, but also by supply, for instance. Still, if we have good estimates of exogenous, idiosyncratic shocks to demand expressed in terms of either prices or output, any of these variables can be treated as the "policy" variable governed by a rule similar to (1.3) with ε_t being the demand shock.

1.3.2 Implementation

The estimation procedure consists of three steps. First, the history of policy shocks is to be estimated. In this chapter I use identified VAR for this purpose for the reasons mentioned earlier.

The next step is to fit a density function $g(x)$ on the estimated policy shocks. It is assumed that they are drawn from the same (zero mean) distribution in each period. Fitting density function can be done either in a parametric or in a nonparametric way. During the simulations I used kernel density estimator with Gaussian kernel and the rule-of-thumb bandwidth proposed by Fan and Gijbels (1996). With the fitted density function one can easily calculate the GPS of any level of the policy variable for any state of the system, because $I_t = \Phi'(Z_{t-1}, Z_{t-2}, ..., e_t) + \varepsilon_t$, and since the variables in Φ' are independent of the current policy shock, $r(I_t, X_t) = r(I_t, \{Z_{t-1}, Z_{t-2}, ..., e_t\}) = g(\varepsilon_t)$.

The final step is estimating the relationship between the outcome and the policy. The explanatory variables are the policy variable and the GPS estimated in the previous two steps. The main advantage of this approach becomes clear at this point: we do not have to control for all the confounders, only for the treatment (policy) and the GPS, and no lags are needed. This is why the GPS-based approach is more promising to identify nonlinear relationship of any form even from few number of observations than the local projection.

The estimation at the third stage follows Hirano and Imbens (2004). With our time series setting, the dependent variable of the regression is the variable of interest (Y) h periods after the observed policy. The explanatory variables are the policy variable (I) and its associated GPS (R). Thus, one have to estimate the following specification by OLS:

$$Y_{t+h} = \beta_{00}^h + \beta_{10}^h I_t + \beta_{01}^h \hat{R}_t + \beta_{20}^h I_t^2 + \beta_{11}^h I_t \hat{R}_t + \beta_{02}^h \hat{R}_t^2 + ... + u_t^h \tag{1.5}$$

Following Hirano and Imbens (2004) I used only second order terms to capture non-linearities. One can, however, include higher orders terms as well. My experience was that higher order terms did not help achieve a better fit.

One issue may emerge when working with macroeconomic variables if the outcome variable is trending while the policy variable not, or vica versa. One example is the estimation of the effect of the central bank's interest rate on GDP. In this case the left hand side variable contains trend, but the right hand side variables not (assuming that the interest rate is stationary). In order to prevent the residual term from absorbing the trend, one can include Y_t as an additional regressor, or to replace Y_{t+h} with $Y_{t+h} - Y_t$. The latter is equivalent to restricting the coefficient of Y_t to be unity in the former.

Our estimate of the average effect of any hypothetical policy i on the outcome can be derived with the following formula (using the second order approximation):

$$\widehat{E[Y_h(i)]} = \frac{1}{T} \sum_{t=1}^{T} \left(\hat{\beta}_{00}^h + \hat{\beta}_{10}^h i + \hat{\beta}_{01}^h \hat{R}(i)_t + \hat{\beta}_{20}^h i^2 + \hat{\beta}_{11}^h i \hat{R}(i)_t + \hat{\beta}_{02}^h \hat{R}(i)_t^2 \right)$$

where $\hat{R}(i)_t$ is the GPS evaluated at the hypothetical policy using covariates at time t.

Finally, the estimated impulse response to a shock sized Δi is

$$\{\hat{Y}_1(i + \Delta i) - \hat{Y}_1(i), \hat{Y}_2(i + \Delta i) - \hat{Y}_2(i), ...\}.$$

1.4 Monte Carlo experiment

Monte Carlo simulations are widely used to compare the sampling properties of various estimators. A common problem is that the choice of the data generating process is necessarily arbitrary to some extent. In the following, I am going to describe my simulation design which aims at delivering a general enough framework in order to obtain robust simulation results. Then the results are shown for three methods, namely, the SVAR, the local projection and the GPS-based approach.

1.4.1 Simulation design

A common practice for making the data generating process realistic is to use a model that is estimated from real data. This is, for example, the approach of Jordà's (2005) first Monte Carlo simulation. This process guarantees that the joint moments of the variables are similar to what is observed in the data, contrary to those simulations generated by a hypothetical model. However, since there are many possible combinations of observable data, the choice of the variables used for the estimation will still contain some arbitrariness.

Instead of relying only on one single data generating process, I will use many, but all based on real data. For this purpose, I collected 22 macro time series for the United States, typically used in empirical studies.[6] All the data are quarterly and cover the period between 1960Q1 and 2006Q4, thus containing 188 observations.[7] In the case of

[6]These are: GDP, consumption, investments, exports, imports, government budget balance, inflation, commodity price index, employment, unemployment rate, hours worked, wage, credit to firms, credit to households, fed funds rate, government bond yields, nominal exchange rate, real exchange rate, stock index, M3, non-borrowed reserves, total reserves. For further details see the appendix.

[7]The only exceptions were exports and imports, with the first observation of the transformed time series being for 1964Q2.

obviously trending variables, such as real GDP, I transformed the time series by taking the first difference of its logarithm. Then I standardized all of them.

The simulation procedure was the following. First, I randomly chose n time series out of the 22, estimated a VAR with a constant, and used these estimates in the data generating process. The number of variables (n) in the VAR was fixed during the whole simulation exercise. The time series of the structural shocks (e_t) were drawn from the standard normal distribution. The elements of the contemporaneous impact matrix (A), which is the collection of the contemporaneous effects of each shock on each variables were also drawn from the standard normal distribution, but $n(n-1)/2$ of them were replaced by zeros. The reason for this is that for estimating A I used the algorithm of Rubio-Ramirez et al (2010), which works only with zero restrictions. Then the time series of VAR residuals (ε_t) were generated according to the relationship $\varepsilon_t = Ae_t$. Finally, simulated series of the n variables were generated using the estimated VAR coefficient matrix and the ε_t-s. The policy variable was also selected randomly out of the n variables with each variable being equally likely chosen.

In the next round I not only drew new random shocks, but I changed the data generating process as well, by fitting a VAR on a new set of n randomly chosen variables and generating a new A matrix, as described above. Thus each simulation was based on a different VAR (each estimated from real data), a different A matrix and different time series of structural shocks.

I conducted three different exercises. In the first one the data generating process was linear, and the SVAR was correctly specified. Naturally, one expects under this environment that the SVAR performs well, but it is still informative to see how the alternative estimators compare to it. The number of lags in the DGP was two, as well as in the SVAR. The local projection was estimated only with linear terms because including higher order terms made the estimates less precise, presumably because of overfitting.

In the second exercise the SVAR model was misspecified. The number of lags in the DGP was based on the Akaike information criterion, but the SVAR model used one lag less. Note that this rule favors the local projection compared to the SVAR, because the latter is forced to be misspecified, while the number lags in the LP regressions were selected by minimizing the Akaike information criterion. Note also that the GPS-based estimator does not contain lags at the final stage (equation (2.3)), thus no lag selection procedure is needed, and thus it cannot be misspecified in this sense.

In the third exercise the DGPs were nonlinear processes. Nonlinearity was modelled as endogenous regime switching. The system had two states, both being a linear SVAR. Regime switching occurred when the sign of the policy variable changed. The contemporaneous impact matrix A did not change across regimes. The impulse responses to be estimated were conditioned on the policy variable being one standard deviation away from its mean[8]. Since the SVAR is linear, its impulse responses cannot be conditioned, thus the unconditional estimates were used. The nonlinear local projection estimator has to be conditioned on all variables, therefore I also calculated the conditional means of all other variables. The size of the shock was 0.25, that is the policy measure under investigation is increasing the policy variable from 1 to 1.25. VAR lag length was selected using the Akaike information criterion, thus, even if the VAR was misspecified due to nonlinearities, the number of lags were determined optimally. The "true" impulse responses were calculated as follows: first a random sample of 10000 observations was generated, then the effect of increasing the policy variable by 0.25 was calculated for each period separately. Finally, the impulse responses starting from observations with the policy variable being between 0.95 and 1.05 were averaged.

In order to obtain pronounced nonlinearities in the third exercise, some model selection were applied. For each draw the mean of the squared differences between the unconditional and the conditional impulse responses were taken across variables in the second quarter. This squared difference is zero for a linear DGP and can be considered as a measure of nonlinearity. At the final stage only the most nonlinear 20 percent of the draws were kept.

In each exercise the SVAR model was estimated first for each simulation using the information about the location of the zero elements in the contemporaneous impact ma-

[8]With standardized variables it practically meant that the policy variable took value of one in the condition.

trix A. This step followed the algorithm of Rubiro-Ramirez et al (2010). This estimate was used to generate the SVAR impulse responses, but also to deliver an estimate of the contemporaneous effect of the policy shock (the appropriate column of $\hat{A}$) for the other two estimators. Moreover, the time series of the policy shocks estimated by the SVAR was also used by the GPS-based estimator. The impulse responses for $h = 1, 2, ...$ were then calculated by the local projection and the GPS-based methods. Since I focus on the point estimates, confidence bands were not calculated.

In the second and third exercises the set of the regressors in the local projection contained squares and cubes of the explanatory variables, as in Jordà (2005). The number of lags were selected by the Akaike information criterion. The GPS-based estimator used only first and second order terms in (2.3), following Hirano and Imbens (2004). Since the local projection and the GPS-based estimator use separate regressions for each horizon h, the estimated impulse responses are typically less smooth than that of the SVAR. This is a strength of the latter. To compensate for that, I considered the smoothed impulse responses of the formers obtained with a nonparametric method: the impulse responses were calculated for 5 quarters longer, and then regressed on the number of quarters of the forecast horizons with Gaussian kernel and the bandwidth selected by the rule-of-thumb approach of Fan and Gijbels (1996). I also experimented with other reasonable kernels and bandwidths, but the qualitative results remained unchanged.

I worked with two sample sizes. The "small sample" consisted of 100 periods. With quarterly data it corresponds to 25 years. The "large sample" consisted of 1000 observations, corresponding to 250 years with quarterly data. In the real life, the small sample is the more relevant case. However, to get an impression about the consistency properties, the large sample results are presented, too.

Figure 1.1 shows an example of estimated impulse responses from one particular data generating process, which is linear, the VAR is correctly specified, and the sample size is 100. On the first three graphs 20 impulse responses from different simulations with the same DGP are presented. The fourth graph shows the means across simulations.

Some features that proved to be quite typical during the simulation can be observed

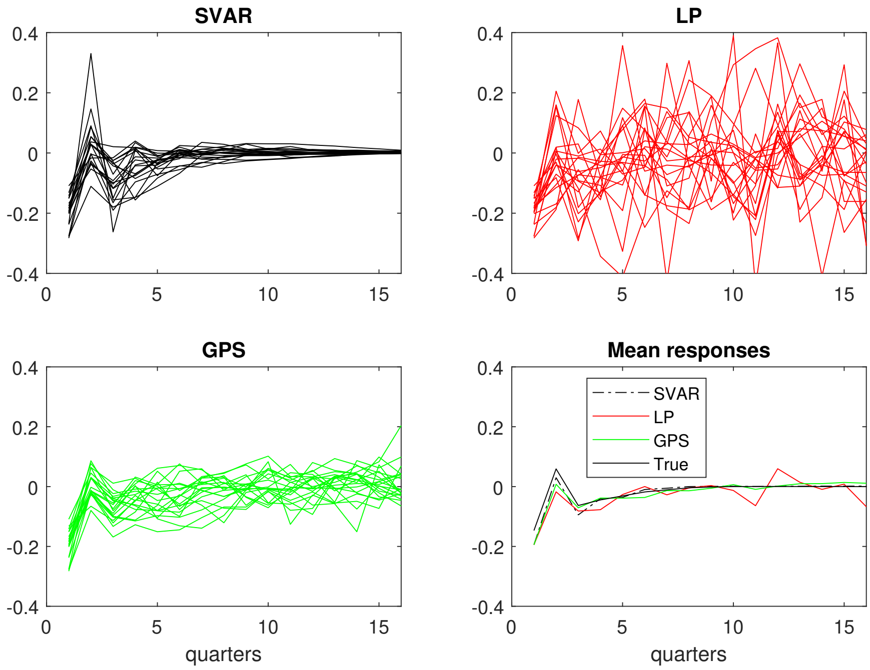

Figure 1.1: Estimated impulse responses from different simulations with the same linear data generating process. Sample size is 100. The VAR is correctly specified.

here. Most importantly, the SVAR estimates are smooth at longer horizons which is a consequence of the iterative way the forecast is made with an autoregressive model. In contrast, the other two estimators give very noisy estimates even beyond 2-3 years. Nevertheless, the mean responses track the true one reasonably well in this particular example.

1.4.2 Results

I evaluated the performance of the three approaches using the root mean squared error. The error was defined as the difference between the true and the estimated impulse response function. The mean of the squared errors was taken not only across simulations, but across the variables as well[9]. Moreover, the quarterly squared errors were transformed to annual mean squared errors by taking simple averages within each year. Thus, for

[9] Recall that all variables were standardized.

example, the overall RMSE for the second year was calculated as the square root of

$$\sum_{h=5}^{8} \sum_{j=1}^{n} (\widehat{IRF}_{j,h} - IRF_{j,h})^2,$$

where $IRF_{j,h}$ is the true, $\widehat{IRF}_{j,h}$ is the estimated impulse response for the jth variable h periods ahead. Since the impulse responses were originally calculated for 16 quarters, the charts show 4 data points (yearly averages) for each estimator.

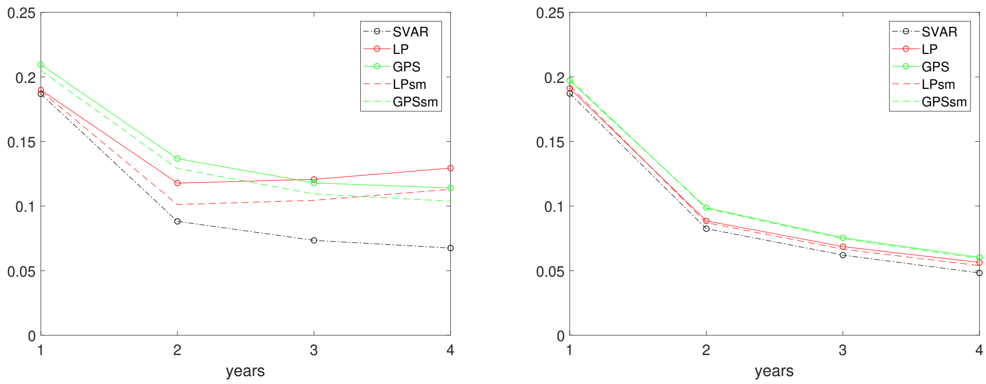

Figure 1.2: Root mean squared errors. 3-variable linear DGP, correctly specified VAR. Left panel: $T = 100$. Right panel: $T = 1000$. "SVAR", "LP" and "GPS" stand for the three estimation methods. Extension "sm" denotes nonparametrically smoothed impulse responses.

For each exercise, results are shown for both sample sizes (100 and 1000), and two different VAR sizes (3 and 6 variables). In the first two exercises 1000 simulations were run to calculate the RMSEs. In the third exercise 10000 simulations were made first.

In some cases, especially with $T = 100$, the estimated VAR was exploding, or fitted the data very poorly. Since the other two methods use outputs of the VAR, their RMSEs became extremely high, too. To ensure that the total RMSE is not dominated by these extreme estimates, I dropped ten percent of the simulations with the highest RMSE for each estimator.

Figure 1.2 shows the RMSEs when the DGP is linear, and the VAR is correctly specified, that is, contains exactly as many lags as were used in the DGP. With a small sample, the SVAR approach outperforms the other two, which is not surprising, because it is a correctly specified model. For one year horizon the other two methods' error is comparable to that of the SVAR, but for longer horizon they yield poorer estimates. The local projection is slightly more precise than the GPS-based estimator for the first two years. When the sample size is increased to 1000, the difference between the SVAR and the other two approaches almost disappears.

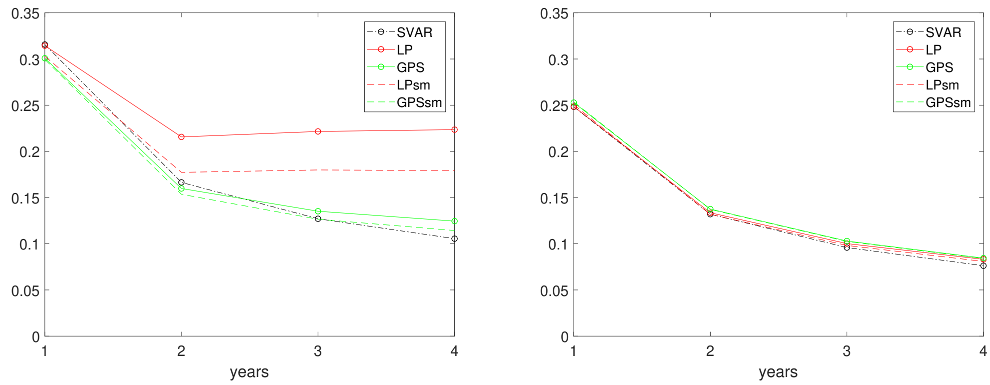

Figure 1.3: Root mean squared errors. 6-variable linear DGP, correctly specified VAR. Left panel: $T = 100$. Right panel: $T = 1000$. "SVAR", "LP" and "GPS" stand for the three estimation methods. Extension "sm" denotes nonparametrically smoothed impulse responses.

With six variables in the data generating process, the overall picture slightly changes as the GPS-based estimator becomes roughly as precise as the SVAR, rendering the local projection the worst estimator (Figure 1.3, left panel). With larger sample size the performances of the three methods are virtually the same, similarly to the 3-variable case (Figure 1.3, right panel).

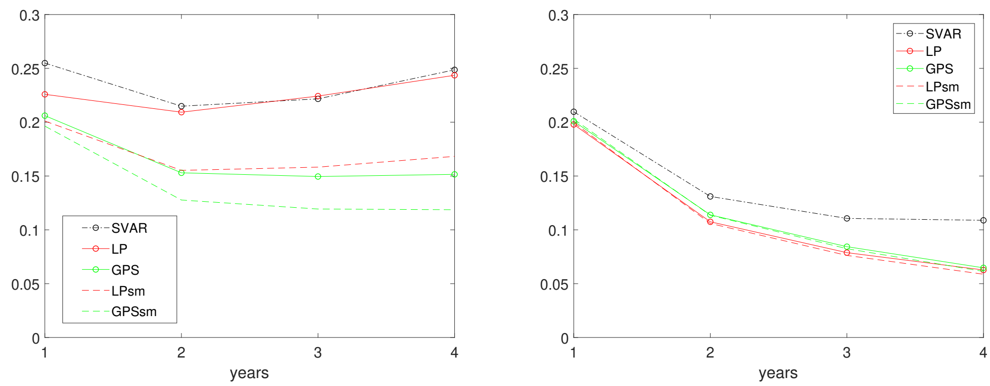

Figure 1.4: Root mean squared errors. 3-variable linear DGP, misspecified VAR. Left panel: $T = 100$. Right panel: $T = 1000$. "SVAR", "LP" and "GPS" stand for the three estimation methods. Extension "sm" denotes nonparametrically smoothed impulse responses.

The lesson from this exercise is that although the performance of the SVAR is obviously the best, using the GPS-based estimator is a better option than the local projection, and in most cases is almost as good as using the SVAR, with the only exception of the 3-variable DGP and small sample.

In the second exercise the VAR is misspecified in the sense that it contains less lags than the DGP. This kind of misspecification may seem to be artificial, because in reality the way people select lag length does not necessarily lead to systematic truncation. On the other hand, there are several algorithms to select the number of lags, therefore, misspecification is very likely. Note also that the other two estimators rely partially on the SVAR estimates, and thus are affected by the same misspecification, even if indirectly. This exercise is informative, because it compares how the competing methods can estimate the impulse response functions when the underlying exogenous shocks come from the same misspecified model.

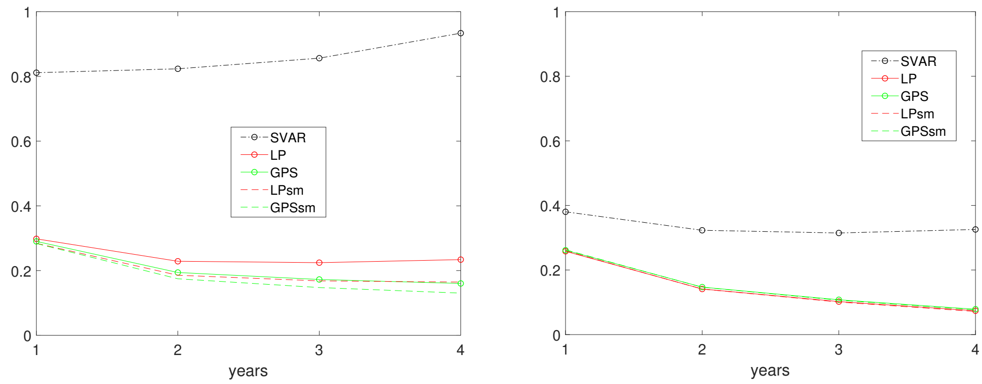

Figure 1.5: Root mean squared errors. 6-variable linear DGP, misspecified VAR. Left panel: $T = 100$. Right panel: $T = 1000$. "SVAR", "LP" and "GPS" stand for the three estimation methods. Extension "sm" denotes nonparametrically smoothed impulse responses.

Figure 1.4 presents the results for the 3-variable case. With a small sample, the GPS-based estimator unambiguously outperforms the SVAR and the local projection at each horizon. The latter is better than the SVAR only with smoothing. The superiority of the GPS-based approach is due to its ability to reduce dimensionality. This seems to be confirmed by the large sample results: while the SVAR remains the worst performing method, the difference between the local projection and the GPS-based estimation disappears.

The ranking of the three methods remains the same when working with 6-variable GPS (Figure 1.5). With the small sample size, the GPS-based method is significantly better than the local projection, again. The only difference compared to the three-variable case is that the relative performance of the SVAR is much worse. The same is true for the large sample case: as previously, the local projection and the GPS-based estimator perform similarly well, while the SVAR produces much larger errors.

Based on the results of the second exercise we can conclude that using the GPS-based method seems to be a very attractive choice when the underlying SVAR model is misspecified.

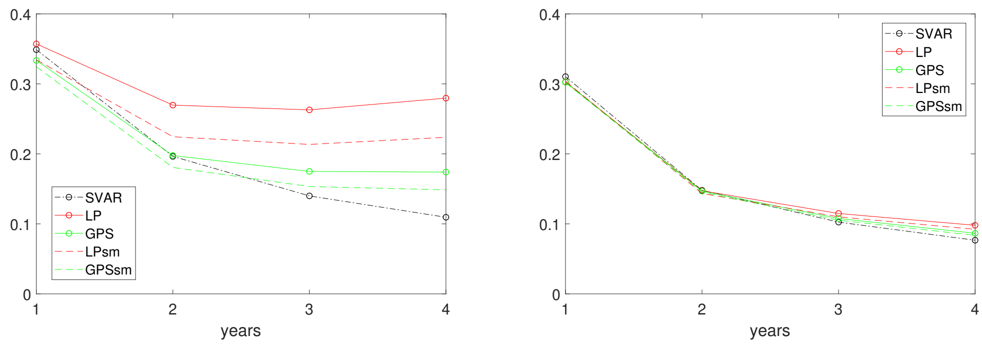

Figure 1.6: Root mean squared errors. 3-variable nonlinear DGP. Left panel: $T = 100$. *Right panel:* $T = 1000$. *"SVAR", "LP" and "GPS" stand for the three estimation methods. Extension "sm" denotes nonparametrically smoothed impulse responses.*

Finally, we compare the performance of our competing methods when the DGP is nonlinear. The SVAR has no chance to capture the nonlinearity, but the other two do have. It should be stressed again that the performance of our nonlinear estimators is influenced negatively by the fact that they use inputs from a misspecified SVAR.

With three variables and a small sample the (smoothed) GPS-based estimator has the smallest RMSE in the first two years (Figure 1.6, left panel). Quite surprisingly, the SVAR does not perform much worse, and for years 3-4 it is the best estimator, despite the fact that it is a linear model. This result can be rationalized by looking at Figure 1.1, because due to the stationarity of the time series, at longer horizons the true impulse response converge to zero, and nonlinearities play negligible role. The local projection produces higher errors than the GPS-approach at each horizon which confirms prior expectations. Interestingly, with larger sample, there is virtually no difference between the three estimators (Figure 1.6, right panel).

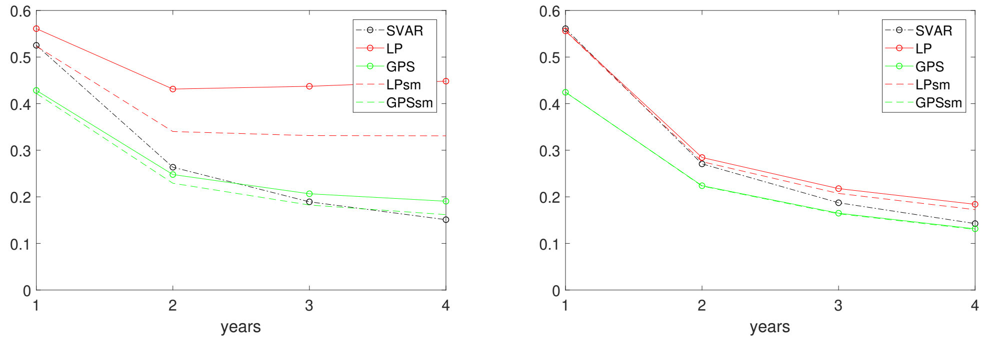

Figure 1.7: Root mean squared errors. 6-variable nonlinear DGP. Left panel: $T = 100$. Right panel: $T = 1000$. "SVAR", "LP" and "GPS" stand for the three estimation methods. Extension "sm" denotes nonparametrically smoothed impulse responses.

With the 6-variable data generating processes, the relative performance of the GPS method improves further as it is clearly the best in the first two years when the sample is small (Figure 1.7). At longer horizons, where the true impulse responses are close to zero and nonlinearities are small, the SVAR performs similarly well. The local projection is the noisiest estimator at each horizon. Increasing the sample size to 1000 the local projection improves a lot, but is still not better than the SVAR. The GPS-based estimator produces the most precise estimates for each horizon considered here.

In sum, the third exercise confirmed that due to its dimensionality reducing property, the propensity score based method performs better than the local projection and, in the first two years, the SVAR. An interesting result is that its advantage does not seem to disappear entirely despite the fact, that the degrees of freedom constraint is less binding.

Based on the results from the three different experiments, one can conclude that the estimator proposed in this chapter is more robust than the SVAR and the local projection, particularly when the sample size is small and the data generating process includes more variables. When the SVAR used for estimating the shocks is correctly specified, it yields almost as good estimates as the two benchmark models. When there is misspecification, the GPS-based estimator's performance is significantly the best, particularly for the first two years.

10.14754/CEU.2021.04

1.5 Conclusion

In this chapter I proposed the use of the generalized propensity score for estimating dynamic causal effect of macroeconomic policy shocks. The GPS is the continuous analogue of the propensity score widely used to estimate treatment effect when the treatment is binary. Allowing treatments to take continuous values is important for macroeconomic policy analysis since most policies can be described properly only by a continuous variable.

Available macro time series are typically short, with low number of observations, making it difficult to identify causal effects, especially when it is nonlinear. Widely used empirical models, such as the SVARs and Jordà's local projection suffer from curse of dimensionality. The generalized propensity score reduces the dimensionality substantially. Once the exogenous shocks to the policy variable are known, the number of parameters to be estimated is much less than with the local projection, which is otherwise flexible enough to capture nonlinearities. This can be very useful when complex dynamic relationship in the data is likely to be present.

A very general Monte Carlo experiment demonstrates that the GPS-based estimator has all the beneficial properties that local projection has, especially its flexibility and robustness to specification choices, contrary to SVARs. Moreover, it performs significantly better than the local projection when the true process is nonlinear and the number of observations is low. In light of the simulation results, the estimator based on the generalized propensity score seems to be the best choice out of the three approaches when there is high uncertainty about the data generating process and the number of observations is low.

Chapter 2

Nonlinear Effects of the Fed's Monetary Policy: Estimation with the Generalized Propensity Score

2.1 Introduction

The causal effect of monetary policy on main macroeconomic variables, such as GDP and inflation has been receiving a distinguished research interest for several decades. The most common approach to estimating this relationship is to use the vector autoregression model (VAR). Exogenous variation in the central bank's interest rate is usually identified with restrictions imposed on the structural decomposition of the VAR residuals, like in Sims (1992), Bernanke and Mihov (1998) or Uhlig (2005). More recently, identification with the help of external instruments has been used (Gertler and Karádi, 2015, Caldara and Herbst, 2019), typically with high-frequency financial data. Although identified or structural VARs (SVARs) have become a convenient tool for such purposes, the linear structure prevents them from estimating nonlinear effects.

Nonlinearity may, however, be an important feature of the monetary transmission mechanism. One argument is that the Phillips-curve is convex, and thus, firms react to increasing demand by raising prices more than they increase output during expansions due to capacity constraints (for example Clark et al, 1996, Boehm and Pandalai-Nayar, 2020). Convexity can be caused by asymmetric price rigidity, too, as in Ball and Mankiw (1994). Schaling (2004) and Dolado et al (2005) discuss the consequences for designing optimal monetary policy.

Another source of nonlinearity can be the financial accelerator, as described in Bernanke and Gertler (1989). During economic downturns firms are more dependent on external financing, the cost of which contains a pro-cyclical premium over the risk-free rate. Since monetary policy affects this external finance premium via its effect on a firm's balance sheet, when the economy is contracting, interest rate changes may have larger impact on financing costs. According to this theory, monetary policy is more powerful during bad times than during expansions.

Non-standard consumer preferences can also lead to nonlinear responses to changes in monetary policy. Santoro et al (2014) presents a model in which households are loss-averse, that is they value the same change in consumption differently depending on its direction with losses having bigger effect on utility than gains. They show that in this case the intertemporal substitution is higher during bad times and, consequently, interest rate changes influence consumption more.

Finally, the role of the zero (or effective) lower bound should be mentioned. The basic idea is that the behaviour of the monetary policy changes when the interest rate is close to zero, which induces nonlinearities in the behaviour of economic agents, too. A straightforward example is that an interest rate cut cannot indicate many further cuts in the future. The nonlinear effects caused by the zero lower bound are analyzed extensively with a New-Keynesian model in Fernández-Villaverde et al. (2015).

In this chapter I investigate whether there are significant nonlinear effects of the Fed's monetary policy on the economy. To address the problem of low degrees of freedom, I use an estimator which is based on the general propensity score (GPS). I adopt the concept of the GPS of Hirano and Imbens (2004) and estimate the response of key macroeconomic variables to an an exogenous monetary policy shock. In the first chapter I demonstrated via Monte Carlo experiments that the GPS-based estimator performs better in small samples than the VAR and Jordà's (2005) local projection in predicting dynamic causal effects when the data generating process exhibits nonlinearities or the VAR is exposed to the risk of misspecification of other form.

Applying the GPS-based estimator to U.S. data, I find that the transmission of the Fed's interest rate decisions depends on the level of the interest rate. When the interest rate is high, GDP and consumer prices react faster and to a greater extent to monetary policy shocks. I find, however, no significant difference between the effect of easing and tightening shocks in absolute terms.

An earlier example for estimating nonlinearities in the monetary transmission mechanism is Weise (1999), who estimated a smooth transition VAR model for the U.S.. He found that monetary shocks have stronger impact on industrial production and weaker on prices when output growth is lower. Moreover, he did not find asymmetry regarding positive and negative shocks.

Tenreyro and Thwaites (2016) estimate a smooth transition local projection model. Local projection has several advantages over VARs, as argued in Jordà (2005), mainly because the joint dynamics of all endogenous variables need not to be modelled in order to estimate the impulse responses. Non-linear effects can also be captured easily even without the smooth transition. According to their results, monetary policy is more powerful during high growth periods. Furthermore, contractionary shocks are found to be more effective than expansionary ones.

Gross and Semmler (2018) investigate how the slope of the Phillips-curve and the effect of monetary policy change with the business cycle in Europe. To estimate the latter, they use a regime switching VAR with the regime changes being driven by the output gap. In line with Tenreyro and Thwaites (2016), they find that the effect of monetary policy shocks is stronger during expansions.

Primiceri (2005) estimates a time-varying VAR for the U.S. economy. Although the results point to significant change in the Fed's reaction function, he estimates no significant change in the propagation of monetary policy shocks.

All the aforementioned approaches suffer from the problem of low degrees of freedom to some extent. Smooth transition models try to reduce the dimensionality by assuming that the nonlinearity is in the form of switching between two regimes. Time varying VARs allow for infinitely many regimes but are unable to say anything about what drives regime changes. Even with the restrictions, these models have at least twice as many parameters to be estimated than the corresponding linear models. Bayesian inference may help in identification, but at the price of further restrictions in the form of priors.

More recently, some new approaches have been proposed to circumvent the problem of low degrees of freedom efficiently. Barnichon and Matthes (2018) employ a basis function approximation of the moving average representation of a VAR. In this way they reduce the number of parameters to be estimated while still being able to identify nonlinear effects. In the empirical application of their study they find that expansionary monetary shocks affect prices more and unemployment less than contractionary ones in the U.S..

Angrist et al (2018) use the concept of the propensity score to estimate dynamic effects of the Fed's interest rate changes. The treatment is the decision of the FOMC at the rate setting meetings. They estimate first a parametric policy reaction function which is used later to derive the propensity score. The average treatment effects are then calculated by inverse probability weighting. Their results show significant differences between the effect of contractionary and expansionary shocks, the former being more powerful with respect to all variables under investigation.

Direct comparison of my findings to those of the above-mentioned papers that investigate how the monetary transmission mechanism is influenced by the cyclical state of the economy, is not possible, because my state variable is the interest rate. However, since the interest rate is typically higher when the economy is in an expansionary phase, my finding, namely, that interest rate changes exert faster and stronger effect on the GDP and prices when the level of interest rate is higher, seems to confirm that of Tenreyro and Thwaites (2016) and Gross and Semmler (2018). Another important result in this chapter is that the sign of the interest rate shock does not cause any nonlinear effects, that is the economy reacts to contractionary and expansionary monetary surprises in the same way in absolute terms. This is in contrast with the findings of Angrist et al (2018) and Barnichon and Matthes (2018).

Being based on the propensity score, the approach employed in this chapter can be considered as a generalization of Angrist et al (2018). Whereas the latter is specific to the discrete nature of the Fed's rate setting decisions, with the use of the generalized propensity score it becomes possible to estimate the effect of policy interventions even in cases when the policy instrument is modelled as a continuous variable.

There are further important differences between the two estimation strategies. The first one is rather technical, namely, the proposed estimator based on the GPS is not a generalization of the inverse probability weighting to the continuous case, but follows Hirano and Imbens' (2004) strategy to estimate dose response functions. The second one is that while with the model of Angrist et al (2018) one can identify only the effect of an interest rate change of a certain magnitude, with the GPS-based approach it is possible to estimate the effect of changing the interest rate from a certain value to another one. That is, the approach in this chapter yields level-dependent estimates by construction. It is a very useful feature if we want to test the hypothesis, for instance, that the monetary transmission becomes weaker above or below a certain threshold level of the interest rate.

The use of the generalized propensity score offers an advantage in the case of small samples over other nonlinear estimators, such as Jordà's (2005) flexible local projection. When estimating the impulse response function, Jordà (2005) includes all the confounders and their lags in the regression. With a nonlinear specification, powers and interactions also appear on the right hand side, and the number of free parameters soon becomes comparable to the number of observations. With the propensity score approach, the full set of controls used only at the first stage, when the policy shocks are estimated by, for instance, a VAR. If the linear approximation is precise enough, there is no need for including higher order terms. At the second stage, when the relationship between the policy and the outcome variables is estimated, the set of regressors collapses to the policy variable, the GPS and their powers and interactions - no lags or other controls are needed to estimate nonlinear effects. Neither the assumption of small number of distinct regimes is necessary.

10.14754/CEU.2021.04

The chapter is structured as follows. Section 2.2 introduces the theoretical background of the propensity score based approach. In section 2.3, the estimation algorithm is described. Section 2.4 presents the result. The final section concludes.

2.2 Estimation based on the GPS

The estimation of the impulse response functions is based on Hirano and Imbens' (2004) concept of the generalized propensity score (GPS), which is the analogue of the propensity score for continuous treatments. They use GPS to identify dose response function from cross-section. In the first chapter I showed how their approach can be adopted to a time series setting. Throughout this chapter I identify the treatment with the level of the central bank's policy rate I_t.

The GPS associated with the treatment status at time t is defined as

$$R_t = f_{i|x}(I_t|X_t), \tag{2.1}$$

where $f_{i|x}(I_t|X_t)$ denotes the conditional density of the interest rate, given the set of covariates X_t. The covariates include lagged values of relevant macro variables and unobserved nonmonetary shocks at period t.

Let us assume an interest rate rule in the following form:

$$I_t = \Phi(X_t) + \varepsilon_t \tag{2.2}$$

where $\Phi(X_t)$ is the expected value of the interest rate, conditioned on the same set of covariates as in (2.1), that is on lagged values of all observed variables, including the interest rate itself, and the (unobserved) contemporaneous shocks other than the monetary policy shock ε_t. If we further assume that the idiosyncratic policy shocks are drawn from the same zero mean distribution in each period, the GPS is simply the density function of this distribution evaluated at ε_t.

Note that contemporaneous observable variables, like GDP, prices etc. are not included in the reaction function in (2.2), which is in contrast with popular interest rate rules, like the Taylor rule (Taylor, 1993). This is because some covariates may react to the interest rate within the same period and thus may not be independent of ε_t. X_t, however, contains only lagged values of observable variables and contemporaneous nonmonetary shocks, that are all independent of the contemporaneous policy shock.

The main challenge in estimating the causal effect of monetary policy is that the central bank sets its interest rate partly in response to the state of the economy, captured by Φ. Thus, macro variables observed at period t and earlier affect not only the outcome (the economy at period $t+1, t+2, ...$), but also the interest rate at period t resulting in biased estimates if the former is directly regressed on the latter. The key condition for removing the endogeneity bias is weak unconfoundedness, which is formulated by Hirano and Imbens (2004) as follows:

$$Y_{t+h}(i) \perp I_t | X_t \text{ for all } i,$$

that is, once we control for all the relevant confounders, potential outcomes $(Y_{t+h}(i); h = 1, 2, ..., H)$ after any hypothetical interest rate (i) are orthogonal to the actual interest rate at period t. It does not mean that the actual outcomes Y_{t+h} are independent from the actual policy at period t, it only says that in the same situation same policies would have had the same effect on the economy, independently of what the actual policy was. As Hirano and Imbens (2004) show, weak unconfoundedness imply that it is enough to control for the generalized propensity score instead of the full set of confounders to eliminate the endogeneity bias.

Macro models usually make the assumption that conditioned on the current state of the economy, the evolution of the endogenous variables in the future is governed by realization of structural shocks in the future. Structural shocks in the future are assumed to be orthogonal to the current state of the economy, including the interest rate at period t, and to have time invariant distribution. All this implies weak unconfoundedness, because once we control for the current state of the economy, future values of the endogenous variables depend only on random variables that are orthogonal to any variable at time t

and earlier, and have a time-invariant joint probability distribution.

As a consequence, the expected outcome will be a function of the policy variable at period t and the associated GPS (R_t):

$$E_t(Y_{t+h}) = f_h^{GPS}(I_t, R_t).$$

Thus, the number of parameters to be estimated is much less than in other models, like Jordà's (2005) local projection, which uses the following model:

$$E_t(Y_{t+h}) = f_h^{LP}(Z_t, Z_{t-1}, ...).$$

where Z_t is the vector of observable confounders, including the policy variable, too. With 6 endogenous variables and 4 lags, for instance, one has to estimate 25 parameters in the linear case, and much more with a nonlinear specification.

Of course, when estimating the policy shocks, that are needed to calculate the propensity score, controlling for many covariates may be needed, because monetary policy reacts to a wide set of information. But this is something that is common with other models that identify the causal effect of the shocks. For modelling the relationship between the policy and the outcomes, only a few number of parameters are to be estimated even with a nonlinear specification. The advantage of using the generalized propensity score materializes at this stage, which is confirmed by Monte Carlo simulations in the first chapter.

2.3 Inference

The estimation procedure consists of three steps (see Chapter 1). First, the history of policy shocks is estimated. Then a probability density function is fitted on the shocks and the generalized propensity scores are derived. Finally, the outcomes are regressed on the policy variable and the GPS, which allows to calculate the effect of any hypothetical policy. Details on the implementation can be found in the appendix.

2.3.1 Data and identification of monetary policy shocks

The exogenous variation of the Fed's target interest rate were estimated with a sign restricted VAR based on Uhlig (2005) for several reasons. First, this is a widely referred study in the topic of estimating the effect of monetary policy. Second, with sign restrictions one obtains more robust results than with just- or overidentified VARs. Finally, the Bayesian framework he used can be easily extended to the GPS-based impulse response analysis, which otherwise would need bootstrapping.

It should be noted that since the monetary shocks are estimated with a SVAR, we approximate the possible nonlinear reaction function Φ in (2.2) with a linear model. It does not mean at all that we assume a completely linear data generating process, because the VAR is not used to estimate the impulse responses. The only assumption is that with the linear approximation of (2.2), we obtain precise enough estimates of the monetary shocks.

Similarly to Uhlig (2005), my sample starts in 1965Q1, but I use ten years more data ending in the last quarter of 2006. The choice of the variables more or less follows the referred paper, too, by using time series of U.S. GDP, GDP deflator, commodity price index, fed funds rate, total reserves and non-borrowed reserves. These time series were taken from the database used for Monte Carlo simulation in the first chapter.[1]

In contrast with Uhlig (2005), I use quarterly data. The advantage of the lower frequency is that the GDP data can be used directly, without interpolation. Another departure from the referred paper is that with the exception of the fed funds rate, I took the first difference of the logarithm of the variables, because the favourable small sample properties of the GPS-based estimator are demonstrated for stationary time series in the first chapter.

The identifying restrictions I used were also different to those imposed in Uhlig (2005) because of the different focus of the study. My aim is to start from a SVAR estimates that produces impulse responses that are common in the literature, first of all a drop in GDP and a gradual decline in consumer prices after an unexpected tightening. Uhlig's (2005)

[1]For further details see the appendix for chapter 1.

main question was whether there is a significant effect of monetary policy on output and his finding was rather mixed. I take the above-mentioned consensus results as given and adjust the identifying restrictions accordingly to investigate whether there is significant deviation from the average effect when different levels of the interest rate are considered.

I assumed that the GDP and the consumer prices do not react to monetary policy shocks within one quarter. I also assumed that tightening shocks have negative effect on the consumer prices one quarter later, and on the non-borrowed reserves within the same and the next quarter. Uhlig (2005) did not imposed zero contemporaneous restrictions on GDP and consumer prices, but assumed negative effect on the latter as well on non-borrowed reserves. In the absence of these zero restrictions, GDP and inflation would jump immediately, which is not consistent with the most typical finding in the literature.

In order to decrease the residual autocorrelation, I included 5 lags in the VAR which is more than what Akaike's and Schwartz's information criteria suggest. According to the multivariate LM-test, even with 5 lags there remained some autocorrelation in the residuals, but increasing the number of lags further did not improve the test result.

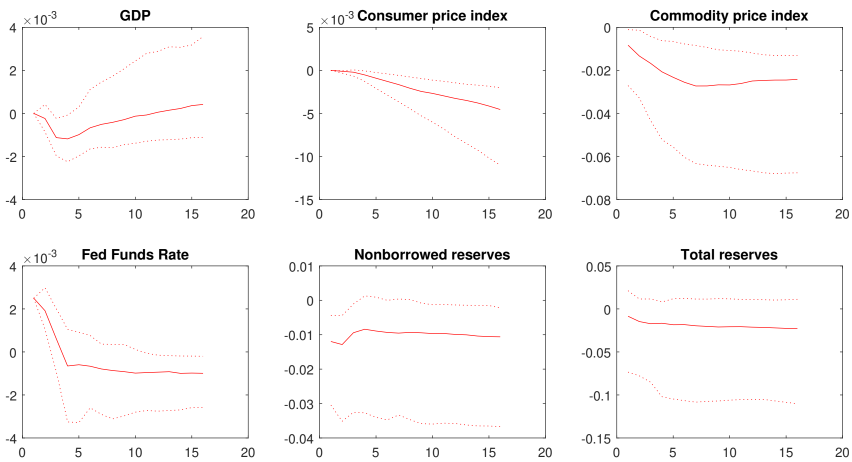

Figure 2.1: Effect of an unexpected 25 basis-point interest rate hike. Estimates with the Uhlig-type SVAR. Solid line is the median, dotted lines are the boundaries of the middle 68 percent of the posterior distribution.

Figure 2.1 shows the estimates of the effect of an unexpected 25 basis-point interest rate hike. The responses are in line with conventional economic thinking and are close to the results of Uhlig (2005). The typical monetary policy shock takes the form of a rather temporary increase in the fed funds rate, lasting only for one year. GDP drops in the first year after the shock, but rebounds in the second. Consumer prices decrease, but very gradually, consistently with rigid price setting. The commodity price index, total and non-borrowed reserves fall permanently.

2.3.2 Deriving the propensity score

The next step is fitting a density function on the estimated policy shocks (ε_t). As mentioned above, it is assumed that they are drawn from the same distribution. I adopted a parametric approach in order to facilitate the Bayesian inference. I assumed normal distribution with zero mean and unknown variance (σ_ε^2). Using the uninformative prior $p(\sigma_\varepsilon^2) \propto 1/\sigma_\varepsilon^2$, the posterior distribution is inverse gamma:

$$\sigma_\varepsilon^2 \sim IG\left(\frac{T}{2}, \frac{\varepsilon'\varepsilon}{2}\right),$$

where T is the number of observations. The GPS for each observation is then computed from the normal density function with zero mean and variance drawn from the distribution above.

2.3.3 Calculating the impulse response functions

The final step is estimating the relationship between the outcome and the treatment. The explanatory variables are the policy variable and the GPS estimated in the previous step. As mentioned earlier, the main advantage of this approach becomes obvious at this point: once we have a good approximation of the policy shocks, we do not have to control for all the confounders, only for the level of the policy variable (treatment) and the GPS, and no lags are needed. This is why this approach may be more capable to identify nonlinear relationship of any form even from few number of observations than, for example, the local projection.

The estimation at the third stage follows Hirano and Imbens (2004). With our time series setting, the dependent variable of the regression (Y) is the variable of interest

(GDP, consumer price index etc.) h periods after the shock. The explanatory variables are the policy variable (I) and its associated GPS (R) at period t. I also included Y_t as a regressor, because some of the outcome variables were highly persistent. The outcome variable Y_{t+h} was approximated with the following model:

$$Y_{t+h} = \alpha^h Y_t + \beta_{00}^h + \beta_{10}^h I_t + \beta_{01}^h \hat{R}_t + \beta_{20}^h I_t^2 + \beta_{11}^h I_t \hat{R}_t + \beta_{02}^h \hat{R}_t^2 + ... + u_t^h, \tag{2.3}$$

where I_t is the interest rate at period t and $\hat{R}_t$ is the associated GPS, the latter calculated in the previous step. I assume that the residual term has a normal distribution with zero mean and variance σ_h^2. Working with the uninformative reference prior again, which is flat for the coefficients and similar to that of the previous step for the variance, the joint posterior distribution of $\beta = (\alpha^h, \beta_{00}^h, \beta_{10}^h, \beta_{01}^h, ...)'$ and σ_h^2 will be a normal-inverse-gamma distribution, that is

$$\sigma_h^2 \sim IG\left(\frac{T-p}{2}, \frac{\hat{u}'\hat{u}}{2}\right),$$

where p is the length of the vector β, $\hat{u}$ is the residuals estimated by OLS, and conditioned on σ_h^2, the posterior of β is multivariate normal:

$$\beta \sim N((X'X)^{-1}X'y^h, (X'X)^{-1}\sigma_h^2),$$

where X is the $T \times p$ matrix of the regressors in (2.3).

Finally, our estimate of the average effect of any hypothetical policy (i) on the outcome can be derived by the following formula (using second order approximation):

$$\hat{Y}_h(i) = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\alpha}^h Y_t + \hat{\beta}_{00}^h + \hat{\beta}_{10}^h i + \hat{\beta}_{01}^h \hat{R}(i)_t + \hat{\beta}_{20}^h i^2 + \hat{\beta}_{11}^h i \hat{R}(i)_t + \hat{\beta}_{02}^h \hat{R}(i)_t^2\right),$$

where $\hat{R}(i)_t$ is the GPS evaluated at the hypothetical policy using covariates at time t.

The estimated impulse response function is then:

$$\{\hat{Y}_1(i+\Delta i) - \hat{Y}_1(i), \hat{Y}_2(i+\Delta i) - \hat{Y}_2(i), ...\}.$$

The posterior distribution of the impulse responses and, thus, the reported error bands reflect both the model uncertainty, stemming from the sign restrictions, and the parameter uncertainty. The latter consists of three components from each stage of the estimation: estimation of the policy shocks, fitting a density function and estimating the relationship between the outcomes and the policy variable. A detailed description of how the posterior distribution of the impulse responses were simulated, see the appendix.

2.4 Results

Figure 2.2 compares the estimated impulse responses of the SVAR model with the unconditional estimates of the GPS-based approach. The shock is a 25 basis-point unexpected rate hike. Since the GPS-based impulse responses depend on the initial level of the interest rate as well as on the size of the shock, I calculated the unconditional effect as the average of the impulse responses conditioned on each interest rate contained by my sample.

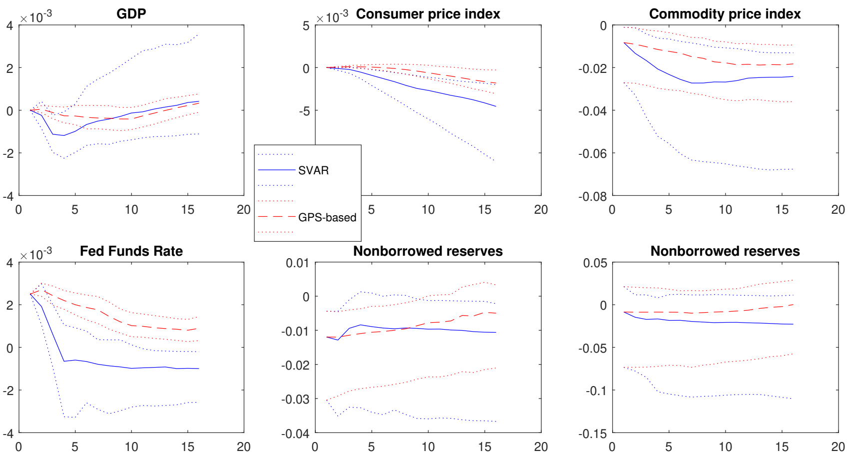

Figure 2.2: Effect of an unexpected 25 basis-point interest rate hike. SVAR model median estimates (solid blue line) and unconditional median estimates with the GPS-based approach (dashed red line). Dotted lines are the boundaries of the middle 68 percent of the posterior distribution.

The point estimates of the contemporaneous effects are the same for both approaches by construction. The GPS-based estimation produces more persistent responses than the SVAR. Nevertheless, the main characteristics are similar. GDP and consumer prices fall after the shock. The same is true for the commodity price index after the initial drop.

Figure 2.3 shows how the effect of the monetary contraction depends on the initial level of the interest rate. I compare the high interest regime with the low one. The high and low interest rates were selected to be the 90 and the 10 percentile of the sample distribution, corresponding to 10.06 and 2.46 percent.

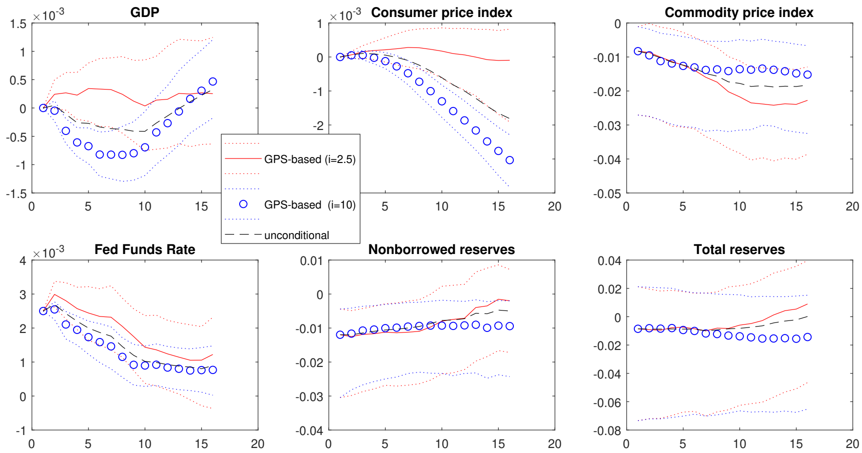

Figure 2.3: Effect of an unexpected 25 basis-point interest rate hike. GPS-based conditional estimates. Solid red line denotes the median of the response when the initial level of the fed funds rate was low, blue circles when it was high. Dotted lines are the boundaries of the middle 68 percent of the posterior distribution.

The most pronounced nonlinearity can be found in the case of the GDP and the consumer prices. Starting from a 10 percent interest rate, a 25 basis-point increase has a faster and stronger effect on both variables, even though the effect on the fed fund rate is roughly the same. This result suggests that monetary policy is more effective when the interest rate is high.

Normally, such periods coincide with high growth periods, and thus, these results seem to confirm those of Tenreyro and Thwaites (2016) and Gross and Semmler (2018), who found that monetary shocks have larger effect during economic expansions. However, the correspondence is limited. In my sample, that is between 1965 and 2006 the correlation between the interest rate and the growth rate of GDP is low, due to the Volcker disinflation period, when the high interest rate were coupled with low growth for a protracted period. Therefore we cannot interpret the results as a direct evidence for a business cycle dependent monetary transmission mechanism. It is true, nevertheless, that the GPS-based estimates would predict more efficient monetary policy for periods characterized by high inflation, high growth and high interest rate.

Regarding the possible nonlinearities with respect to the direction of the interest rate change, Angrist et al (2018) and Tenreyro and Thwaites (2016) found that rate hikes have larger impact on the economy than rate cuts. The results shown in Figure 2.3 cannot be directly related to them, as the correlation between the level and the first difference of the interest rate is virtually zero in my sample. Still, we can say that the GPS-based estimates suggest that if the economy is in an expansionary phase and the interest rate is higher than average, monetary tightening is a powerful countercyclical measure, and, similarly, when the economy is in depression (and the interest rate is low), interest rate cuts are less effective.

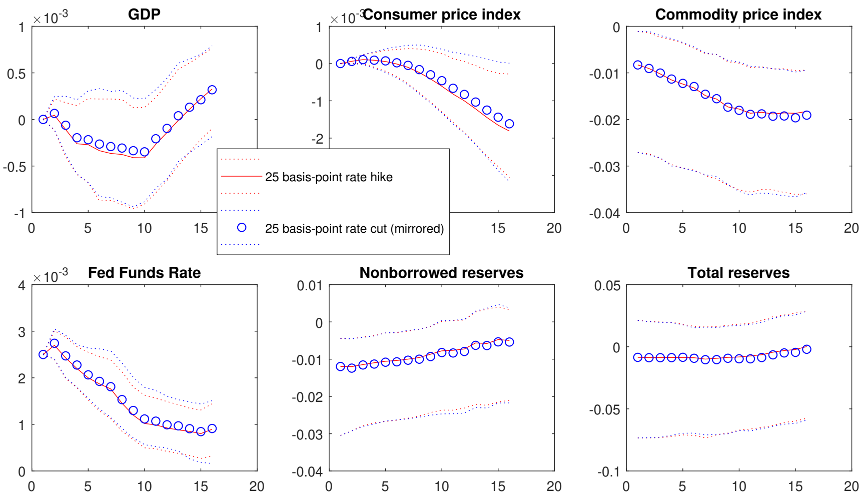

Figure 2.4: Effect of an unexpected 25 basis-point interest rate change. GPS-based estimates conditioned only on the size of the shock. Solid red line denotes the median of the response to a rate hike, blue circles to a rate cut, multiplied by minus one. Dotted lines are the boundaries of the middle 68 percent of the posterior distribution.

However, the possibly asymmetric effect of monetary tightening and easing can be estimated directly with the generalized propensity score. Since it is the change of the interest rate we are interested in, the estimates are unconditional with respect to the level of the interest rate and are calculated in a similar manner as described in the beginning of this section. Figure 2.4 compares the effect of a 25 basis-point rate hike and a rate cut of the same magnitude, with the responses to the latter being multiplied by minus one. Clearly, there is no significant difference, which is in contrast with the corresponding findings of Angrist et al (2018), Barnichon and Matthes (2018) and Tenreyro and Thwaites (2016), but in line with that of Weise (1999).

2.5 Conclusion

In this chapter the estimator proposed in the previous chapter was applied to U.S. data in order to investigate whether the effect of the Fed's interest rate decisions depend on the initial level of the interest rate or on the size of the rate change. In Chapter 1 I demonstrated that estimation based on the generalized propensity score may give reliable estimates of possibly nonlinear, dynamic causal effects even if the number of observation is low.

Although the exogenous variation in the interest rate were estimated with a linear SVAR model, the results displayed significant nonlinearities. Most importantly, interest rate decisions seem to exert faster and larger impact on GDP and consumer prices when the interest rate is high. This is more or less in line with earlier findings, although only with the additional assumptions that the interest rate is positively correlated with the business cycle.

On the other hand, dependence on the size and the direction of the interest rate change does not show signs of nonlinearity. This finding is in contrast with those of earlier estimates in the literature.

Chapter 3

Estimating the Effect of Monetary Policy with Dissenting Votes as Instrument

3.1 Introduction

There is a large body of empirical literature on the effect of monetary policy. Most of the studies use identified or structural VAR to estimate this effect. The main challenge is to circumvent the endogeneity problem, namely that the economy and the monetary policy may react to the same shock rendering causal interpretation of impulse responses impossible. The most common approach is to take the linear combination of the VAR residuals that meets some predefined identifying restrictions, and to call this time series exogenous policy shocks, that is the source of the exogenous changes in the policy variable.

Recently, Stock and Watson (2012) proposed using external instruments for identifying exogenous shocks. External here means that the exogenous variation of the policy variable does not come from the VAR residuals, but from an additional variable not included in the VAR. Mertens and O. Ravn (2013) used external instrument in a SVAR to estimate the effect of tax changes. Gertler and Karádi (2015) estimated the effect of monetary policy using the surprise content of the interest rate decisions as external instrument.

In this chapter I use voting records of the Fed for constructing an instrument for the policy rate. There are two features of central banks' decision making process that may generate exogenous variation in monetary policy. The first is rounding: the policy rates are in most cases multiples of one quarter of a percentage point. The second feature is the consensus-seeking principle, that is the interest rate supported by as many members as possible is chosen as the decision of the committee. Riboni and Ruge-Murcia (2014) call the these features "size friction" and "decision-making friction", respectively. In another study (Riboni and Ruge-Murcia, 2010) they consider different voting protocols and find that the consensus model fits major central banks' interest rate decisions best. Due to the above-mentioned frictions, the decision outcome may differ from its expected value, and part of this difference can be regarded as an exogenous policy shock.

A simple example could be the case when roughly half of the decision making body finds current level of interest rate appropriate, and the other half would like to cut it by 25 basis points. If individual members' preferences are more or less (but not perfectly) known by the public, the expected value of the decision will be a 12.5 basis point cut. The final decision may depend on one or two votes in an almost random way. Either holding the policy rate or cutting by 25 basis points would hit the market as a surprise (either 12.5 or -12.5 basis points) and can be considered as exogenous variation in the sense that it is not correlated with economic variables that influence both the policy rate and the future path of the economy.

Dissenting votes indicate the direction of the surprise. If the final decision is a cut, but dissenting members would have held the interest rate, the policy shock is an easing shock. In the opposite case when the decision making body does not change the policy rate but several members would have done it, the outcome is a tighter than expected monetary policy.

In line with the tendency to make monetary policy more transparent, many central banks publish voting records of committee members. I use the history of FOMC members' dissent record collected by Thornton and Wheelock (2014). I create an index of dissent from this dataset to approximate the sign and size of the monetary policy shock for each

10.14754/CEU.2021.04

rate setting decision. Then I estimate a VAR from data on the main U.S. macroeconomic variables and instrument the policy rate residual of the VAR by the index.

Within the same proxy-SVAR framework, Gertler and Karádi (2015) build on the surprise movements in Fed funds futures around the announcements of FOMC decisions. As emphasised in Jarociński and Karádi (2020), this instrument contains not only policy shocks, but information shocks, too. The latter comes from the Fed's assessment of the state of the economy which may differ from market participants' view. If the central bank has more precise information on trends in the economy, the surprises in the interest rates may be dominated by that knowledge, which is correlated with state of the economy. Consequently, the estimated effects are biased. Jarociński and Karádi (2020) disentangle information shocks from monetary shocks by imposing additional identifying assumptions, and show that the effect of monetary policy estimated by them is different from what Gertler and Karádi (2015) found.

The same problem is less likely to arise with the dissent instrument proposed in this chapter. Although the distribution of votes indicate where the interest rate would be in the absence of decision-making frictions, and this depends on the central bank's assessment of the state of the economy, both the actual and the "frictionless" rate reflect the same assessment, and thus their difference is not much related to the possible information asymmetry between the central bank and the public. The proposed measure of dissent is a proxy for that difference, therefore, when using as external instrument in a SVAR, no additional identifying restrictions are necessary to eliminate the information shocks.

I find that monetary policy shocks have significant effect on real variables with the expected sign, especially in the long run. An unexpected tightening lowers GDP, industrial production and employment. On the other hand, the response of nominal variables are insignificant, and their sign contradicts to the conventional views of the transmission of monetary policy and is reminiscent of the "price puzzle" phenomenon.

These qualitative results remain robust when I change the variables in the VAR, the frequency of the time series and the sample. The conventional explanation of the price puzzle (Sims, 1992) is based on the argument that central banks look at expected future inflation when setting interest rates and use information not included in the econometrician's data. Castelnuovo and Surico (2010) argues that the omission of forward-looking variables generates price puzzle especially when the monetary policy does not respond strongly enough to inflation, just as in the pre-Volcker era. With the dissent-based identification, however, changing the sample period or including variables such as inflation expectations and commodity prices does not alter the big picture. Therefore, as long as the dissent index is a valid instrument, the price puzzle seems to be a feature rather than an artifact.

The chapter is structured as follows. First, I present simulations with a stylized model of collective decision making and demonstrate that dissents can be a valid instrument for identifying exogenous variation in the policy variable. Then I introduce an index of dissents created from FOMC members voting record. In section 3.4 the methodology of the estimation is described. In section 3.5 the proxy-SVAR results are discussed. Section 3.6 concludes.

3.2 Simulation with a stylized model of collective decision making with dissent

Making group decision is a more complex procedure than individual decision-making if there are diverse views on the optimal outcome. Monetary policy committees typically make strong effort to reconcile views of its members in order to arrive at a decision that has strong enough support to convince the public that the policy change will be persistent. The consequence of this effort is that the final outcome will almost never coincide with any member's most preferred policy.

Riboni and Ruge-Murcia (2014) call the desire to reconcile different views on the optimal interest rate "decision-making friction". One tool to decrease diversity among members is to restrict the available options, which typically means considering only interest rate changes that are multiple of 25 basis points. Riboni and Ruge-Murcia (2014) call this "size-friction".

Gerlach-Kristen (2004) finds that dissenting votes can help predict future interest rate decisions of the Bank of England. Riboni and Ruge-Murcia (2014) demonstrates that this result is due to the two aforementioned frictions. In an earlier paper (Riboni and Ruge-Murcia, 2014) they show that major central banks make interest rate decisions in a way that is observationally equivalently to the consensus-seeking model, which implies these frictions.

In the following I will demonstrate how information on dissenting votes may help solve the endogeneity problem. First, I highlight the main intuition with the help of some very simple examples. Then I present simulations with a stylized model of collective decision-making characterised by frictions to show that the measure of dissent can be a valid instrument.

3.2.1 The main intuition

First consider a simple example, in which the decision making body consists of a single member. Of course, in this case there is no friction due to consensus-seeking, but it can shed light on how the "optimal" decision is distorted by the size-friction, that is by restricting the set of options to some discrete values.

Let us assume that the decision maker can choose only 0 or 1. Also assume that the optimal value of the policy instrument is a random number j drawn from the $U(0,1)$ uniform distribution. As she cannot choose numbers between 0 and 1, the policy maker rounds the optimal value to the nearest integer J. Clearly, the expected value of J is 0.5, and thus the surprise component of the decision is $J - 0.5$, which is positively and strongly correlated with the distortion, that is with $J - j$, because they always have the same sign.

Now let us add another "friction" to the decision making procedure, and consider the case when the decision making body consists of three members. Each member's preferred value is drawn from the same $U(0,1)$ distribution, independently from each other. Then they all choose either 0 or 1 depending on which one is closer to the individually preferred policy. The final decision is 0 when at least two members chose it, and 1 otherwise. Due to symmetry, the expected value of the decision-making body's choice is 0.5 again. Thus, the surprise is -0.5 with probability 0.5, and 0.5 with the same probability.

With more than one member dissent can occur. In this example there are dissenting votes if two members vote for 0 and one for 1, or two votes for 1 and one for 0. One measure of dissent can be the number of dissenting votes for lower value minus the number of dissenting votes for higher value divided by the number of total votes. In our case this measure can take the values of $-1/3, 0$ or $1/3$.

Since our measure of dissent is either zero (unanimous voting) or has the same sign as the surprise component of the decision, there is positive comovement between the two. It is very straightforward to derive that the standard deviation of the dissent variable and the surprise is $1/(2\sqrt{3})$ and $1/2$, respectively, their covariance is $1/8$, and thus their correlation is $\sqrt{3}/2 = 0.866$.

These very simple examples showed the main intuition behind looking at dissenting votes as a potential proxy for shocks generated by the decision-making frictions. In real life monetary policy decisions are related to the underlying economic developments that can influence disagreement in a systematic way. If this is the case, a measure of dissenting votes may not be valid instrument and cause bias in estimation.

To see how the state of the economy may interfere with collective monetary policy decision-making, let us suppose that individual preferences regarding the policy instrument are formed by not only pure random shocks but they are state-dependent, too. Let the individually preferred policies be

$$j_i = x + w_i$$

where x is either 0.25 or 0.75 (”low” or ”high” states) with 0.5-0.5 probabilities, and the idiosyncratic policy shock w_i is drawn from the uniform distribution $U(-0.5, 0.5)$ for each member independently. Otherwise the model is the same as previously. Consequently, in the low state the distribution of the individually preferred value of the policy variable is $U(-0.25, 0.75)$, while it is $U(0.25, 1.25)$ in the high state. The possible policy outcomes are still 0 or 1, but their conditional probabilities are different, as Table 3.1 shows. Note that the expected value of the decision is 5/32 in the low state and 27/32 in the high state.

$x = 0.25$				$x = 0.75$			
outcome	dissent	surprise	cond. prob.	outcome	dissent	surprise	cond. prob.
0	0	-10/64	27/64	0	0	-54/64	1/64
0	1/3	-10/64	27/64	0	1/3	-54/64	9/64
1	-1/3	54/64	9/64	1	-1/3	10/64	27/64
1	0	54/64	1/64	1	0	10/64	27/64

Table 3.1: Possible outcomes, the corresponding surprises and conditional probabilities in the two states.

In the low state, the probability that the dissenting vote is higher than the common decision (second row of Table 3.1) is higher (27/64) than the opposite case (9/64), in which the dissenting member votes for a lower rate (third row of Table 3.1). In the high state the opposite is true. This implies correlation between the direction of dissent and the economic shock. Indeed, it can be shown that the correlation between the measure of dissent (defined as previously) and the state of the economy is now 0.375. Nevertheless, the dissent still exhibits stronger comovement with the shocks not related to the underlying economic developments. Assuming that the expectations are formed conditionally on the state of the economy, the correlation between the surprise caused by the decision and the measure of dissent is 0.678.

Although the information content of the vote distribution may be partly driven by the underlying economic shocks, this is not necessarily true always. To see this, consider a modification of the previous example. The only change is that the state variable x can now take the values of 0.5 and 1.5 with the same probabilities. It implies that the distribution of the member specific preferences is $U(0, 1)$ in the low state and $U(1, 2)$ in the high state. The possible outcomes are 0 or 1 in the low state and 1 or 2 in the high state with each conditional probability being 0.5.

Note, that in this case the conditional distribution of the dissent measure is completely the same in the two states (and the same as in the second example), therefore, there is no comovement between the state and our measure of dissent at all, while the correlation between the latter and the surprise is 0.866, again. The explanation of this is that now the effect of the economy on the decision is large enough compared to the distance between the available options.

Based on these simple examples we can conclude that the distribution of dissenting votes is generally correlated with the difference between the expectations and the final outcome. Whether it is correlated with potential confounders depends on the relative size of the shocks and the distance between the potential outcomes of the decision. To give an impression about these dependencies, a simulation with a more detailed model is shown in the next subsection.

3.2.2 The model

In this model[1] the state of the economy is captured by an autoregressive process, and the policy makers vote according to a time-varying rule which is heterogeneous across them. Since we are only interested in whether a proper measure of dissent can be a proxy of exogenous shocks to the policy, there is no feedback from the policy variable to the economy.

The state variable x_t follows an AR(1) process with mean 0:

$$x_t = \rho_x x_{t-1} + u_t^x$$

The decision making body consists of n members. They make decision on the policy variable j in each period. Member i's preferred level of the policy variable is described by the policy rule:

$$j_{it} = \phi_{it} x_t + v_t + w_{it}$$

where w_{it} is a white noise idiosyncratic shock to member i's ($i = 1, ..., n$) decision at period t, v_t is a white noise common policy shock. Idiosyncratic shocks capture unpredictable deviation by a member from her policy rule that are unrelated to others' decision. Common shocks can occur, for example, because the committee's decision is based partly

[1] The model shares many features with those presented in Riboni and Ruge-Murcia (2014) and Gerlach-Kristen (2008), but there are also important differences due to the different purpose.

on the staff's economic analysis, and thus any mistake made by the staff can result in unexpected synchronized deviations from the individual policy rules.

The time-varying member-specific reaction coefficient ϕ_{it} is another AR(1) process with mean 1:

$$\phi_{it} = 1 + \rho_\phi(\phi_{i,t-1} - 1) + u^\phi_{ti}.$$

The innovations $(u^x_t, u^\phi_{ti}, v_t, w_{it})$ are white noise processes uncorrelated with each other.

The collective decision is based on individually preferred outcomes according to the following rule: first, each member rounds her preferred level of the policy variable to the nearest integer (J_{it}). Second, the integer number J_t receiving the most votes is chosen as the body's decision. If the mode is not unique, one of the most popular outcomes is picked randomly with equal probabilities for each.[2]

Clearly, the difference between the final decision and the mean of individual preferences comes from two sources: rounding and the collective decision rule, namely that the mode, not the mean of the individual votes is the collective choice. Each of these effects can be considered random as long as the variance of x_t and/or of the policy shocks is large enough to generate many potential decision outcomes.

To evaluate the validity of a potential instrument, we need a definition of the policy shock, which is not straightforward. Generally, the deviation of the policy variable from the policy rule is meant by it. In our model the systematic behaviour of the decision making body cannot be easily captured by a single feedback rule.

Another approach identifies policy shocks with the difference between the actual decision and the decision expected by economic agents. The two approaches are identical if the public is aware of both the common policy rule and the state of the economy, and forms expectations rationally.

Throughout the chapter I define the policy shock as the difference between the body's

[2] In real life the last round of the procedure is more complex, since the final outcome typically has to enjoy the support of the (qualified) majority of the members. If the mode has no majority, further rounds of voting take place until a (qualified) majority is formed. In the simulation we consider a simpler decision rule by taking the mode as the final decision. However, the same intuition behind the proposed instrumental variable remains valid even if a majority rule is applied, as in Riboni and Ruge-Murcia (2014).

final decision and its expected value,

$$u_t^j = J_t - E(J_t), \tag{3.1}$$

where E denotes the model consistent expectation with the information set including the state of the economy (x_t) and the policy rule of each member (the ϕ_{it}s), but not the common and the individual policy shocks.

Let us denote the mean of the expected values of the individually preferred policy by $\tilde{E}(J_t)$, that is

$$\tilde{E}(J_t) = \frac{1}{n}\sum_{i=1}^{n} E(j_{it}) = x_t \frac{1}{n}\sum_{i=1}^{n}\phi_{it}.$$

The policy shock then can be decomposed into five terms as follows:

$$u_t^j = J_t - \tilde{E}(J_t) - \left(\tilde{E}(J_t) - E(J_t)\right) = J_t - \frac{1}{n}\sum_{i=1}^{n}\phi_{it}x_t - \left(\tilde{E}(J_t) - E(J_t)\right) =$$

$$= J_t - \frac{1}{n}\sum_{i=1}^{n}(j_{it} - v_t - w_{it}) - \left(\tilde{E}(J_t) - E(J_t)\right) =$$

$$= J_t - \frac{1}{n}\sum_{i=1}^{n} j_{it} + \frac{1}{n}\sum_{i=1}^{n} w_{it} + v_t - \left(\tilde{E}(J_t) - E(J_t)\right) =$$

$$= \left[J_t - \frac{1}{n}\sum_{i=1}^{n} J_{it}\right] + \frac{1}{n}\sum_{i=1}^{n}(J_{it} - j_{it}) + \frac{1}{n}\sum_{i=1}^{n} w_{it} + v_t - \left(\tilde{E}(J_t) - E(J_t)\right)$$

Each component has an intuitive interpretation: the first one is the distortion caused by the aggregation rule (picking the mode instead of mean), the second is the distortion caused by rounding, the third is the average of the idiosyncratic shocks, the fourth is the common decision shock. The last term is the difference between the model consistent expected value and the mean of the individual expected values, which is, according to the simulation results, small in magnitude[3]. Consequently, the final outcome differs from the expectations not only because of the unexpected decision shocks, but also due to the

[3] To highlight the difference between the two "expected values", consider a simple example with a three-member decision making body. Each member votes for 0 with 1/3, and for 1 with 2/3 probability. Clearly, the mean of the individual expected values is 2/3. The model consistent expected value which takes the decision making mechanism into account is $3 * (2/3)^2 * (1/3) + (2/3)^3 = 20/27$, which is slightly more than 2/3.

nature of the decision making mechanism, namely, rounding and selecting the mode.

It should be noted that this is not an orthogonal decomposition. While the third and fourth terms are orthogonal to each other by construction, the same is not necessarily true for the first two and for their relationships with the decision shocks.

The first term, the deviation of the collective decision from the mean of the individual votes can be observed as long as the individual votes are disclosed. My proposed proxy for the monetary policy shock is thus the average magnitude of dissents made public by several central banks. To be a valid instrument, it should be correlated with the policy outcome (J_t), but uncorrelated with other shocks in the economy, which is u_t^x in our model. The first condition is likely to be met as long as the other four terms do not completely offset its correlation with the policy shock. The second condition can be met if there is enough variation in the shocks (at least in their effect on the policy variable) compared to the distance between possible outcomes, which is typically 25 basis point in the case of central banks' policy rate.

In the empirical application to be presented later, I use the record of dissenting votes of the Fed collected by Daniel L. Thornton and David C. Wheelock, in which only the direction of the dissent, not the exact value of the alternative votes is indicated. Thus, following Riboni and Ruge-Murcia (2014), I consider the net balance of dissenting votes created as the proportion of votes for weaker action minus the proportion of votes for stronger action, or equivalently

$$d_t = \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(J_{it} < J_t) - \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(J_{it} > J_t), \tag{3.2}$$

where $\mathbb{I}$ is the indicator function taking value of one when the expression in parenthesis is true, and zero otherwise. Because of the signs in the definition, this proxy is expected to be correlated positively with the policy shock.

It should be stressed that this measure of dissent captures the effect of decision-making frictions on the collective decision rather than the degree of disagreement among members. To see the difference, note that, for example, in a 9-member committee the proxy is zero when there is no dissent, similarly to the case when two members voted for higher and two members for lower interest rate. Thus, it is not directly related to the uncertainty, which influences the second moment of the voting distribution mainly, only indirectly.

3.2.3 Simulation results

Since the relevance and the validity of the proposed instrument cannot be derived analytically, not even with this simple model, numerical simulations are needed to see the statistical properties of voting balance and how it is related to the underlying shocks.

During the simulations I kept the autoregressive parameters fixed: $\rho_x = \rho_\phi = 0.9$, that is the shocks hitting the economy and the policy preferences have fairly persistent effect. As for the other parameters, I experimented with numerous combinations. The standard deviation of v_t, w_{it}, x_t and ϕ_{it} changed independently from 0.05 to 0.25 with 0.05 steps. This means 625 combinations of the four parameters. The decision making body consisted of 7 members ($n = 7$). For each parameter combination I made simulation for 10,000 periods.

To calculate the model consistent expectations some approximation was needed, because the shocks are unbounded and, consequently, the number of possible individual outcomes is infinite even after rounding. As a first step, I simulated the individual votes for a given parameter combination for each period, and assumed that the public knows the distribution of votes from this simulation and considers only potential decision outcomes between the first percentile minus one and the 99th percentile plus one. The true probabilities of the outcomes outside this range were added to those of the two extreme outcomes. Since these probabilities are very small, this kind of truncation of the true distribution results only in negligible distortion when calculating the model consistent expected value.

The policy shock u_t^j is then calculated as the difference between the collective decision and this expectation. It should be noted that this model consistent shock is not equal to the one defined for illustration purpose in (3.1), but according to the simulation results, they are very close to each other.

The validity of the instrument requires zero correlation with u_t^x and non-zero correlation with J_t. Figure 3.1 shows the joint empirical distribution of these correlation coefficients.

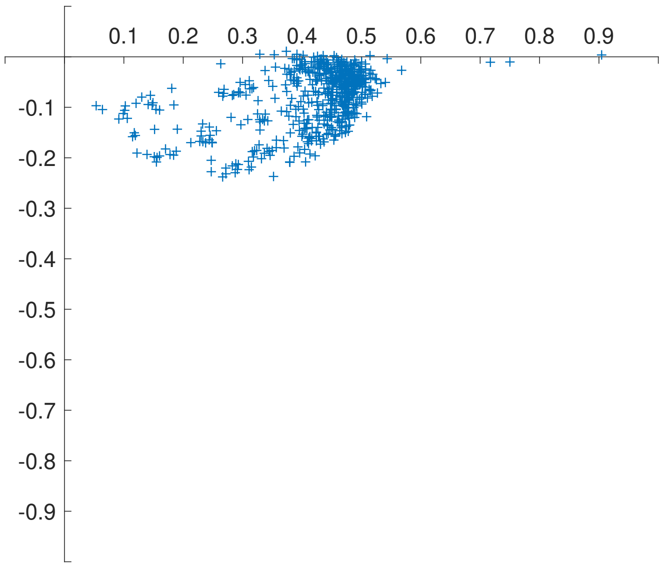

Figure 3.1: Simulated correlations of the dissent proxy with the policy variable (horizontal axis) and the economic shocks (vertical axis).

With most parameter combinations the dissent index variable defined in (3.2) proved to be valid with the exception of very few cases when the correlation with the economic shocks was significantly nonzero (negative). It should be noted here that increasing further (above 0.25) the standard deviation of the policy shocks, the preference shocks and the economic shocks would result in even more valid instrument.[4] In the very few cases when either the correlation with policy shocks was low or the correlation with economic shocks was far from zero, the variance of the shocks in the model were rather low. It is thus worth further investigation, in what parameter regions will our proposed instrument be invalid.

One benchmark for choosing the empirically relevant parameter combinations is the variability of the dissent index. The standard deviation of the monthly instrumental variable used in the empirical application was 0.087, and that of the quarterly one was

[4]For the same reason as explained in the previous footnote.

0.155. Thus, I consider simulation parameters producing similar statistics, that is standard deviation of the proxy variable between 0.08 and 0.16.

Within the empirically relevant parameter region defined above, the correlation between the proxy and the policy outcome is still high. The mean is 0.46 and 90 percent of the distribution is above 0.27. The correlations with the economic shocks are close to zero but with a fat tale in the negative territory. The mean of the distribution is -0.083, the 10th and 90th percentiles are -0.174 and -0.016, respectively. Thus, even if in most cases the instrumental variable can be considered as valid, there is a non-negligible part of the empirically relevant parameter region in which the instrument co-moves with the economic shocks.

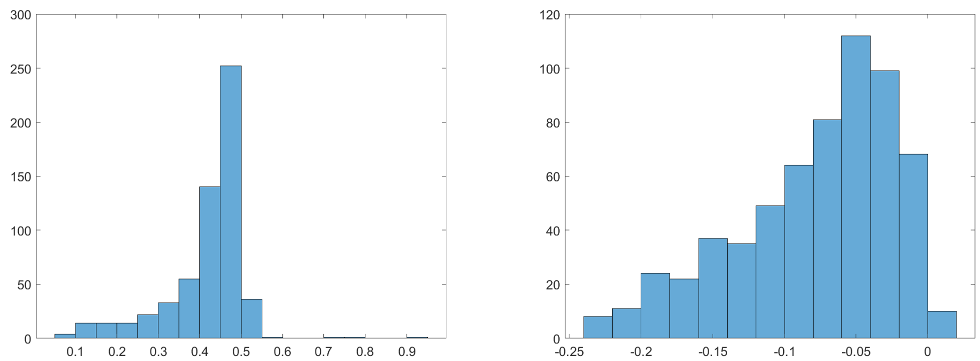

Figure 3.2: Histograms of simulated correlations with the policy variable (left panel) and the economic shocks (right panel) when the standard deviation of the dissent proxy is in the empirically relevant region.

Table 3.2 summarizes the main statistics of the two parts of the empirically relevant parameter region, where in the second part the correlation between the dissent index and the economic shock is less than -0.1, containing 36 percent of all empirically relevant parameter combinations. Since the mean of the reaction coefficients (ϕ) is one, and both the individual and the common policy shocks enter the reaction function directly, the standard deviation of the economic shocks and the two types of policy shocks can be compared directly in terms of their contribution to the variance of the policy variable.

The main difference between the two parameter regions is that in the first one, when the instrument is valid as defined above, the variance of the common decision shocks is significantly larger than that of the individual decision shocks. When the correlation

	standard deviation of				correlation of d_t with		
	x_t	v_t	w_{it}	ϕ_{it}	J_t	u_t^j	u_t^x
valid	0.1622	0.186	0.116	0.15	0.442	0.484	-0.045
invalid	0.155	0.095	0.19	0.155	0.357	0.429	-0.153

Table 3.2: Means of standard deviations and correlations with the proxy variable in the two parameter regions where the instrument can be considered valid (first row) and invalid (second row). x_t: state of the economy, v_t: common decision shocks, w_{it}: individual decision shocks, ϕ_{it}: individual reaction function coefficients, J_t: policy variable, u_t^j: total policy shocks, u_t^x: economic shocks.

between the instrument and the economic shocks is less than -0.1, this relation turns to the opposite as the standard deviation of the individual shocks becomes twice as large as that of the common ones. Nevertheless, the instrument seems to be relevant in each case, as the average correlation with the policy variable is above 0.35.

Unfortunately, in real life we cannot directly observe the relative variances of all shocks incorporated in this simple model. Thus, we cannot *a priori* decide whether the proposed variable will be a valid instrument in empirical applications. All what we can conclude from this simulation exercise is that it will less likely be valid when the common decision shocks are significantly smaller than the individual ones. The main intuition behind this result is that disagreement can generally be driven by the state of the economy, but with large enough common decision shocks it becomes unpredictable in which direction the final decision will be diverted from the optimal policy by the decision-making frictions.

3.3 The instrumental variable

To approximate the distance between the collective outcome and the mean of individually preferred outcomes, I use the record of dissents on FOMC monetary policy votes, which is an extended version of the database constructed by Daniel L. Thornton and David C. Wheelock and used in their study, Thornton and Wheelock (2014). This database contains the voting records with dissents for all FOMC meetings since 1936.

The information from this database I use is the number of members with dissenting votes for tighter and easier policy actions. For the period when the FOMC targeted the federal funds rate, the exact distance between the common decision and the average of individual targets would be a good proxy for the exogenous policy shock described above. In the absence of numerical record of each member's preferred interest rate, I use the net balance of tighter and easier preferences, normalized by the total number of voters. That is, my instrumental variable is

$$d_t = -\frac{n_t^+ - n_t^-}{N_t}, \tag{3.3}$$

where n_t^+ and n_t^- is the number of votes for tighter and easier policy action, N_t is the number of total votes. The minus sign serves only normalization purposes, because in this way we can expect the the instrument to have positive correlation with the policy shock caused by the collective decision making mechanism. Dividing by the number of votes is motivated by the time-varying size of the decision making body. Apart from the minus sign at the beginning, this index is exactly the same as the first one Riboni and Ruge-Murcia (2014) use for forecasting future interest rate decisions.

For the months without FOMC meeting, the value of this variable is set to zero. When working with quarterly data, I aggregated the dissents by taking the sum of monthly data in each quarter. The monthly evolution of the instrumental variable is presented in Figure 3.3.

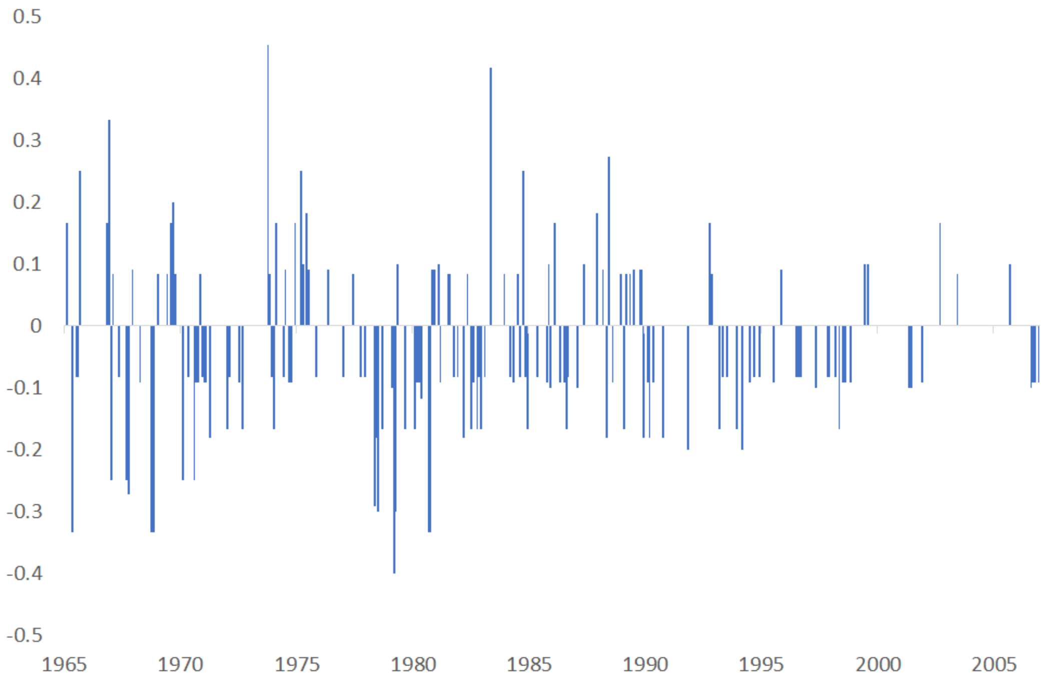

Figure 3.3: Net balance of dissenting votes as a ratio of total votes.

I also consider alternative proxies based on the assumption that the policy shocks can be larger when the Fed's interest rate target is changing, or when the previous decision caused a big surprise. During prolonged periods of predictably unchanged interest rate, the information content of dissenting votes can be smaller, since the alternative interest rate target they suggest has smaller probability as an outcome. Conversely, when the Fed surprises the market with an interest rate change, the length of the cycle and the level of the fed funds rate in the medium term is more uncertain, thus the surprise content of the next decisions can be larger. Accordingly, the distribution of votes can convey more information about what other outcomes might have been plausible.

According to the argument above, I define alternative instruments, too. To overweight dissents during times of rapidly changing interest rates, I multiply d_t by the absolute change of the effective fed funds rate in the previous period or by the absolute value of the lagged residual of the VAR's interest rate equation. As it turns out, the alternative variables are significantly better instruments than the basic one introduced in (3.3).

3.4 Methodology

The proxy-SVAR estimation relies on the assumption that the instrumental variable is correlated with the policy shock, but not with the other structural shocks. With a well specified VAR model, the contemporaneous impact of the policy shock on the VAR's endogenous variables can be consistently estimated by regressing the VAR's non-policy residuals on the policy residuals with the proxy variable as an instrument.

Let x denote one non-policy variable, Y the vector of all endogenous variables. The equation of the VAR corresponding to x can be written as (ignoring exogenous observables and the intercept, and assuming only one lag)

$$x_t = a^x Y_{t-1} + \varepsilon_t^x,$$

where a^x is the corresponding row of the coefficient matrix and ε_t^x is the residual term.

Similarly, the equation of the policy variable i is

$$i_t = a^i Y_{t-1} + \varepsilon_t^i.$$

The residuals are linear combinations of the unobservable structural shocks, with one of them being the policy shock denoted by e_t^p. Particularly, let

$$\varepsilon_t^x = s_{np}^x e_t^{np} + s_p^x e_t^p$$

and

$$\varepsilon_t^i = s_{np}^i e_t^{np} + s_p^i e_t^p,$$

where e_t^{np} is the vector of non-policy shocks, s_{np}^x and s_{np}^i are the corresponding weight vectors (the impact responses), s_p^x and s_p^i are the contemporaneous effects of the policy shock on x and i, respectively.

If d_t is correlated with e_t^p but not with the other structural shocks, then regressing ε_t^x on ε_t^i with the proxy variable d_t as instrument, the regression coefficient will asymptotically be equal to s_p^x/s_p^i. To calculate the contemporaneous effect on each variable, one has to use the fact that $E(SS') = \Sigma$ where S is the matrix of contemporaneous effect of all structural shocks on all endogenous variables and Σ is the variance-covariance matrix of reduced form residuals. For further details see footnote 4 in Gertler and Karádi (2015).

Because some of my instrumental variables are only weakly correlated with the residuals of the interest rate equation, for inference I use the approach of Montiel Olea et al (2020), which is robust to instrument weakness.

3.5 Results

My benchmark VAR consists of monthly observations of fed funds rate (as the policy variable), employment, PCE deflator, commodity price index, non-borrowed reserves and

M2 money aggregate, just as in Jordà (2005)[5]. The fed funds rate is monthly average, The instrument is the interaction of the dissent index and the absolute value of the previous period's rate change, measured by the first difference of the fed funds rate variable. As mentioned earlier, the interaction of the pure dissent index with some measure of previous period's interest rate surprise results in stronger correlation with the contemporaneous interest rate.

The sample starts in January 1985 and ends in December 2006. All variables, except from the fed funds rate, are log-differenced. The charts, however, present the impulse responses in level. The lag length is 6.

The choice of this particular specification, sample and instrument was motivated by the relatively low autocorrelation in the residuals according to the LM-test, significant cross-variable effects according to the Wald-test, and high F-statistics in the first stage regression. The latter was 7.64, which is still lower than the rule-of-thumb threshold of 10 recommended in Stock et al (2002). This is why the weak-instrument robust inference of Montiel Olea et al (2020) was employed. It should be noted, however, that conventional inference with the plug-in estimator and δ-method confidence sets would lead to the same qualitative conclusions.

Figure 3.4 presents the impulse responses to a 25 basis point rate hike shock. The higher interest rate prevails for a relatively long, more than one year period. Employment decreases, as expected, and reaches its trough four years after the shock. A bit surprisingly, employment does not seem to recover: even 8 years after the shock it is still half percent lower than initially, and this difference is statistically significant even at 5 percent.

The behavior of the PCE deflator is different from what standard theories of monetary transmission mechanism predict, as prices rise gradually, despite the subdued activity of

[5]The data used in this chapter are from the same sources as those of the first chapter. Additional monthly data used only in this chapter are from the FRED database: "All Employees: Total Nonfarm Payrolls, Thousands of Persons, Monthly, Seasonally Adjusted", "Private Consumption Expenditure Deflator: All Items Non-Food Non-Energy for the United States (DISCONTINUED), Index 2010=1, Monthly, Seasonally Adjusted", "M2 Money Stock, Billions of Dollars, Monthly, Seasonally Adjusted", "Industrial Production Index, Index 2012=100, Monthly, Seasonally Adjusted", "Consumer Price Index for All Urban Consumers: All Items in U.S. City Average, Index 1982-1984=100, Monthly, Seasonally Adjusted".

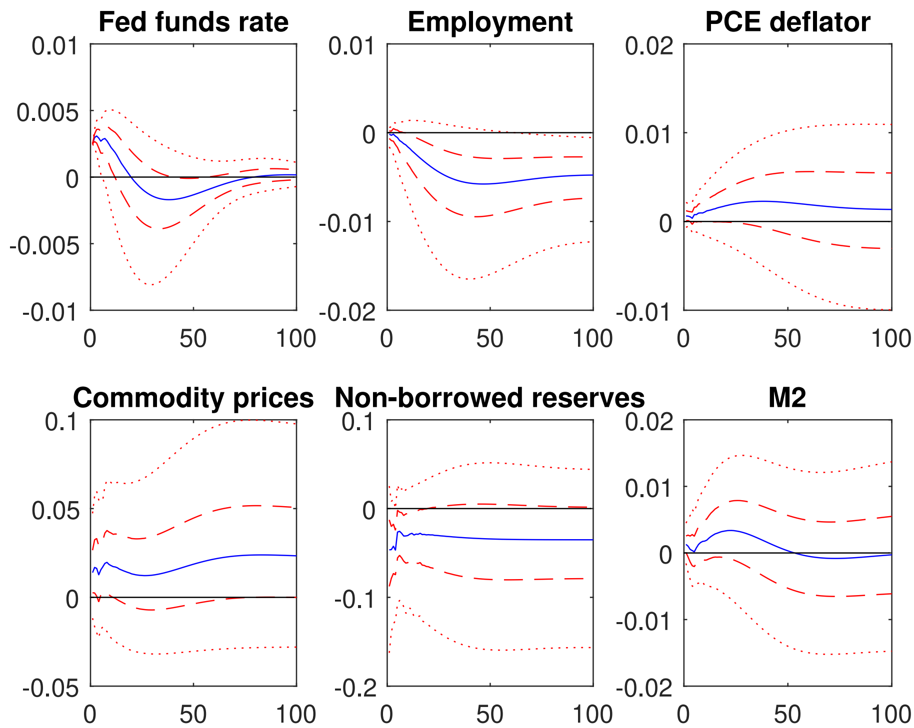

Figure 3.4: Effect of an unexpected 25 basis point interest rate hike. Estimation from monthly data between January 1985 and December 2006. Point estimates as well as 68 percent and 95 percent confidence intervals are calculated with the code used in Montiel Olea et al (2020).

the real economy. Nevertheless, the positive effect is not significant statistically at the conventional levels. The same is true for commodity prices, but with an immediate jump. Non-borrowed reserves drop immediately, M2 rises temporarily, with both responses being statistically insignificant.

In the appendix results from alternative VARs are shown. The specification choices were motivated by low residual autocorrelation, joint significance of the VAR coefficients corresponding to variables other than own lags, and the F-test of instrument strength, just as in the benchmark case. In each case the VAR consisted of one real variable (employment, industrial production or GDP), one measure of overall price level (PCE deflator or CPI), and some other controls (inflation expectations[6], commodity price index, M2, non-borrowed reserves, total reserved, and Dow Jones Industrial Average index).

The findings from benchmark estimation are quite robust to changing the specifica-

[6]I used the logarithm of the ratio of the mean level forecast for 5 and 4 quarters ahead from the Survey of Professional Forecasters

tion, including the variables in the VAR, lag length, sample period, data frequency and instrument. The variable capturing the real activity drops gradually, but persistently and quite significantly. The overall price level increases, but this increase is insignificant in most cases.

This pattern is different from conventional view of the monetary transmission mechanism, which predicts falling prices and output (and employment) after a contractionary shock. It is rather reminiscent of the price puzzle, found in many empirical studies investigating the effect of monetary policy [7]

There are several explanations for the price puzzle. One branch of the arguments attributes it to identification failure. Sims (1992) argues that the Fed uses more information than a typical VAR contains. He demonstrates that the inclusion of commodity prices mitigates the puzzle, because it conveys extra information about future inflation. Castelnuovo and Surico (2010) goes further in this direction and show that the omission of inflation expectation is a problem only in the pre-Volcker period, that is when the estimation sample ends before 1979. Their explanation is that prior to Volcker's chairmanship, the Fed's reaction to the inflation was weak, generating sunspot shocks to inflation expectations. Without controlling for them, monetary policy shock estimates may be distorted. They also show that widely used identification schemes do not produce significant price puzzle for the post-Volcker period.

Another branch of the literature argues that increasing prices after a monetary contraction are not statistical artifact, but rather the genuine response of the economy. The explanation of Barth and Ramey (2002) is based on the cost channel. When interest rates are higher, financing working capital becomes more expensive, thus a monetary contraction causes a negative supply side shock with falling output and increasing prices. They emphasise, however, that this channel influences the total effect of monetary policy only in the short run.

An entirely different explanation of the price puzzle is offered by the neo-Fisherian theory (Uribe, 2018). It is based on the Fisher-equation which establishes a positive

[7] For a detailed literature survey see Rusnak et al (2013).

relationship between inflation and interest rate. Because monetary policy cannot influence the real interest rate in the long run, if the nominal interest rate is changed permanently, it is the inflation rate that has to adjust in the same direction in the long run. This effect exists even in New-Keynesian models (Garín et al, 2018).

Finally, the Fiscal Theory of Price Level also predicts rising prices after a monetary tightening in the case of active fiscal policy rule (Sims, 2011). The main intuition is that after an interest rate hike, the present value of future budget surpluses, which determines the real value of government bonds, decreases, and the equilibrium on the bond market restores through increasing consumer prices.

The results presented in this chapter can be considered as another evidence for increasing prices after a monetary contraction as long as the identifying restriction, namely, that the distribution of dissenting votes are uncorrelated with economic shocks is reasonable. An important feature of my results is that the price puzzle is present even if inflation expectations or commodity prices are included, and also for the post-Volcker period, as results from the alternative specifications show.

Uribe (2018) estimates the effect of short run and long run changes in the interest rate separately. He finds that while a temporary rate hike decreases both output and inflation (just as in the conventional view on the monetary transmission mechanism), permanent rate hikes increase both inflation and output, consistently with the neo-Fisherian effect.

My results can be reconciled with Uribe's (2018) result. Since I do not distinguish between permanent and transition interest rate shocks, the monetary shocks identified in this chapter are presumably a mixture of them. This is supported by the fact that the fed funds rate's reaction to the initial monetary shock is more persistent than in Uribe (2018) after a temporary shock. If temporary shocks are frequent enough, linear combination of output and price level responses of Uribe's (2018) empirical model can easily produce falling output and rising prices after a "mixed" monetary shock, just as the results presented here.

10.14754/CEU.2021.04

3.6 Conclusion

In this chapter I presented a novel way to identify the effect of monetary policy on the economy. I used record of dissenting votes on the FOMC rate setting meetings as an instrument. Due to rounding and majority voting the final outcome is likely to differ from the expected outcome. I used a stylized model to demonstrate that this difference is generally unrelated to the underlying economic shocks.

With this instrumental variable, I estimated several proxy-SVARs for the U.S. A very robust finding from the most relevant specifications is that monetary tightening depress the real economy permanently. Another important finding is that consumer prices increase after the monetary contraction, although the impulse responses are less significant than those of the real variables.

The results presented in this chapter can be considered as another evidence for price increase after a monetary contraction, a phenomenon known as price puzzle in the literature. Although there are some arguments that the puzzle is a statistical artifact due to identification failure, my approach avoids these traps to some extent. A possible theoretical explanation for the puzzle found in this chapter is that unexpected interest rate changes are sometimes permanent ones, that have the opposite effect on inflation, in line with the neo-Fisherian theory.

Appendix A

Appendix for chapter 1: Data sources

Time series from the FRED database:

Short name	Definition (long name)	Transformation
CRDQUSAHABIS	Total Credit to Households and Non-Profit Institutions Serving Households, Adjusted for Breaks, for United States, Billions of US Dollars, Quarterly, Not Seasonally Adjusted	deflated by Consumer Price Index: Total All Items for the United States, Index 2010=100, Quarterly, Seasonally Adjusted (CPALTT01USQ661S), log difference
EXPGSC1	Real Exports of Goods and Services, Billions of Chained 2012 Dollars, Quarterly, Seasonally Adjusted Annual Rate	log difference
FEDFUNDS	Effective Federal Funds Rate, Percent, Monthly, Not Seasonally Adjusted	quarterly average
FGOSNTQ027S	Federal government; operating surplus, net, Flow, Millions of Dollars, Quarterly, Seasonally Adjusted Annual Rate	deflated by Consumer Price Index: Total All Items for the United States, Index 2010=100, Quarterly, Seasonally Adjusted (CPALTT01USQ661S)

Short name	Definition (long name)	Transformation
GDPC1	Real Gross Domestic Product, Billions of Chained 2012 Dollars, Quarterly, Seasonally Adjusted Annual Rate	log difference
GDPDEF	Gross Domestic Product: Implicit Price Deflator, Index 2012=100, Quarterly, Seasonally Adjusted	log difference
GPDIC1	Real Gross Private Domestic Investment, Billions of Chained 2012 Dollars, Quarterly, Seasonally Adjusted Annual Rate	log difference
IMPGSC1	Real imports of goods and services, Billions of Chained 2012 Dollars, Quarterly, Seasonally Adjusted Annual Rate	log difference
IRLTLT01USQ156N	Long-Term Government Bond Yields: 10-year: Main (Including Benchmark) for the United States, Percent, Quarterly, Not Seasonally Adjusted	none
LRHUTTTTUSQ 156S	Harmonized Unemployment Rate: Total: All Persons for the United States, Percent, Quarterly, Seasonally Adjusted	none
MABMM301USQ 189S	M3 for the United States, National Currency, Quarterly, Seasonally Adjusted	deflated by Consumer Price Index: Total All Items for the United States, Index 2010=100, Quarterly, Seasonally Adjusted (CPALTT01USQ661S), log difference
NNUSBIS	Narrow Effective Exchange Rate for United States, Index 2010=100, Monthly, Not Seasonally Adjusted	first difference of quarterly average of logarithm
NONBORTAF	Non-Borrowed Reserves of Depository Institutions Plus Term Auction Credit (DISCONTINUED), Billions of Dollars, Monthly, Seasonally Adjusted	first difference of quarterly average of logarithm

10.14754/CEU.2021.04

Short name	Definition (long name)	Transformation
PCECC96	Real Personal Consumption Expenditures, Billions of Chained 2012 Dollars, Quarterly, Seasonally Adjusted Annual Rate	log difference
PRS85006013	Nonfarm Business Sector: Employment, Index 2012=100, Quarterly, Seasonally Adjusted	log difference
PRS85006022	Nonfarm Business Sector: Average Weekly Hours, Percent Change at Annual Rate, Quarterly, Seasonally Adjusted	none
QUSNAMUSDA	Total Credit to Non-Financial Corporations, Adjusted for Breaks, for United States, Billions of US Dollars, Quarterly, Not Seasonally Adjusted	deflated by Consumer Price Index: Total All Items for the United States, Index 2010=100, Quarterly, Seasonally Adjusted (CPALTT01USQ661S), log difference
RESBALNS	Total Reserve Balances Maintained with Federal Reserve Banks, Billions of Dollars, Monthly, Not Seasonally Adjusted	first difference of quarterly average of logarithm
RNUSBIS	Real Narrow Effective Exchange Rate for United States, Index 2010=100, Monthly, Not Seasonally Adjusted	first difference of quarterly average of logarithm
USAHOUREAQISMEI	Hourly Earnings: Manufacturing for the United States, Index 2010=100, Quarterly, Seasonally Adjusted	log difference

Time series from other sources:

Name	Transformation	Source
CRB Commodity Price Index	first difference of quarterly average of logarithm	Coibion (2012)
Dow Jones Industrial Average, monthly, end of period close value	first difference of quarterly average of logarithm	S&P Dow Jones Indices LLC, a division of S&P Global

Appendix B

Appendix for chapter 2: Implementation details

The calculation of the posterior distribution of the impulse response functions consisted of the following steps.

1. Draw from the VAR parameters' posterior, as in Uhlig (2005).

2. Generate a randomly drawn candidate shock vector, i.e. the vector of contemporaneous effect of the monetary policy shock on the six endogenous variables. This step also follows Uhlig (2005), but with the first two entries, corresponding to GDP and consumer prices, being zero, according to our identifying restriction.

3. Check, whether the sign restrictions are satisfied with the VAR parameter and shock vector draws. If yes, draw 10 times from the posterior of the policy shocks' variance.

4. For each shock variance draw, draw from the posterior of the parameters of regression (2.3).

5. Calculate the impulse response functions.

6. Repeat from step 2, but maximum 10 times.

7. Repeat from step 1, until 1000 draws collected.

10.14754/CEU.2021.04

In the regression model (2.3) I used first and second order terms, similarly to Hirano and Imbens (2004). Since the coefficient estimates of terms including the GPS were often insignificant, but higher powers proved to be significant, I also experimented with a version in which terms up to seven order were the regressors. The results became more noisy, but did not affect the main findings qualitatively.

Unconditional impulse responses to a 25 basis-point rate hike were estimated in the following way: the impulse response were calculated for 10 interest rate levels randomly drawn from the sample with replacement in each round, and then the mean of these impulses responses were taken. Unconditional here means that the expectation is not conditioned on the interest rate, only on the size of the shock.

Appendix C

Appendix for chapter 3: Results from alternative VAR specifications

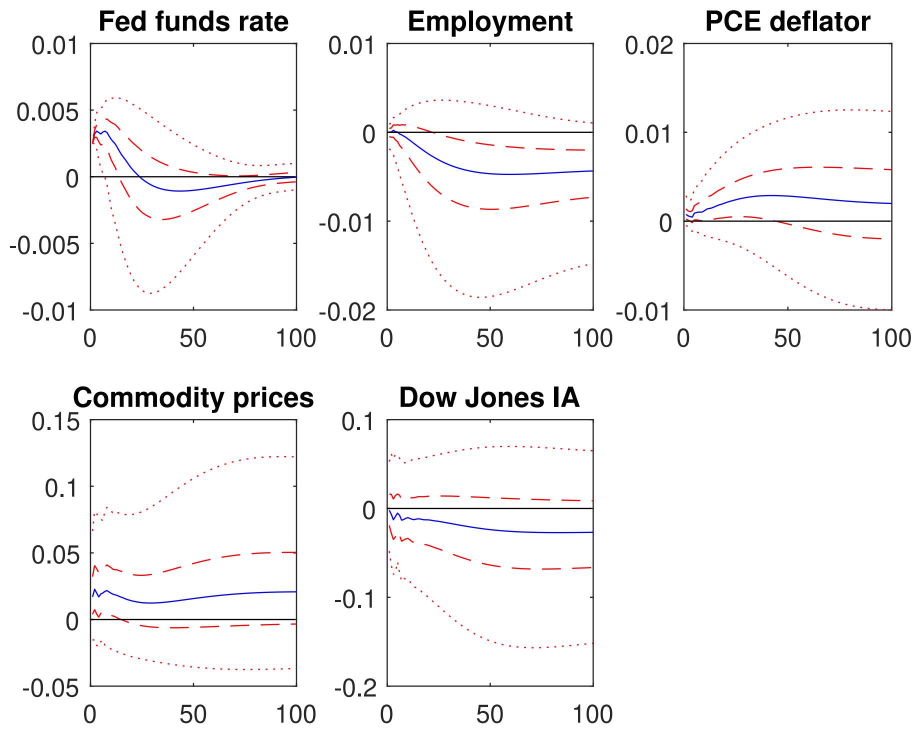

Figure C.1: Effect of an unexpected 25 basis point interest rate hike. Estimation from monthly data between January 1985 and December 2006. VAR includes 6 lags. The instrument is the product of the dissent index and the fed funds rate's change in the previous month. The first stage F-statistics is 7.27. Point estimates as well as 68 percent and 95 percent confidence intervals are calculated with the code used in Montiel Olea et al. (2020).

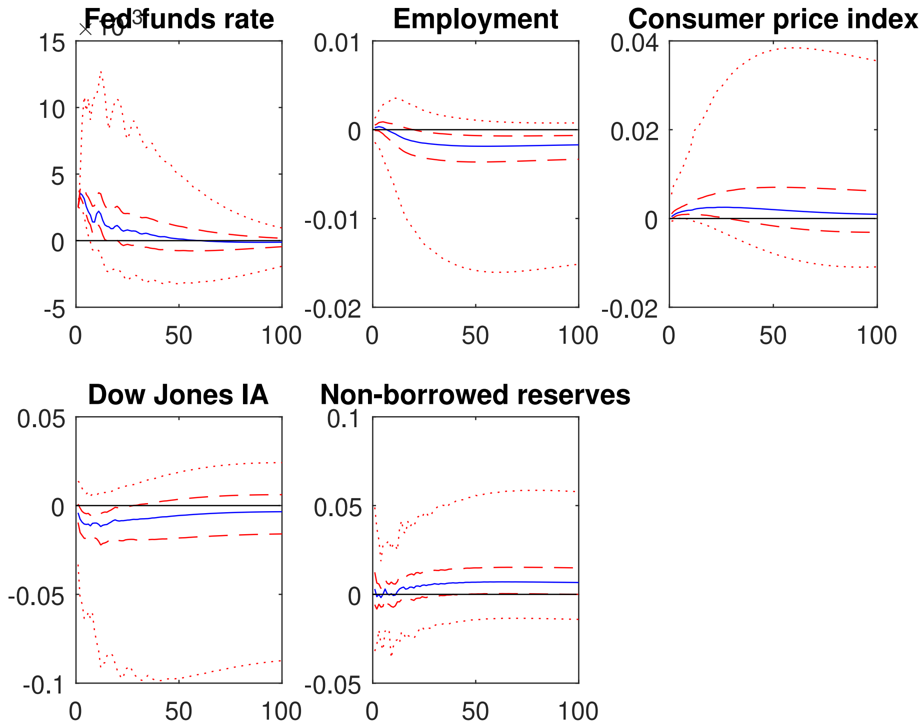

Figure C.2: Effect of an unexpected 25 basis point interest rate hike. Estimation from monthly data between January 1971 and December 2006. VAR includes 12 lags. The instrument is the product of the dissent index and the lagged estimated residual in the interest rate equation of the VAR. The first stage F-statistics is 8.5. Point estimates as well as 68 percent and 95 percent confidence intervals are calculated with the code used in Montiel Olea et al. (2020).

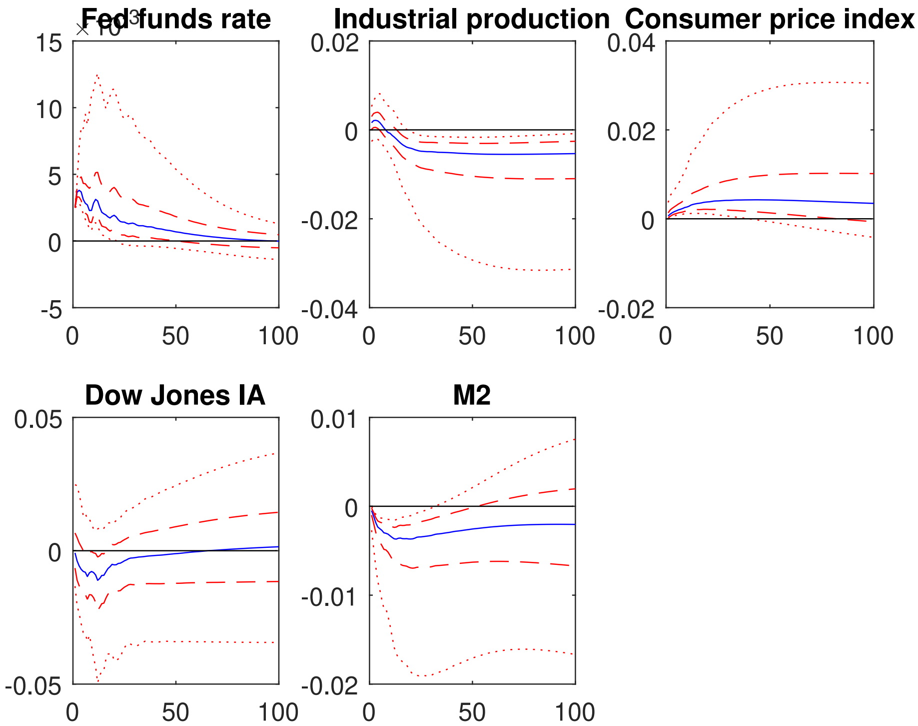

Figure C.3: Effect of an unexpected 25 basis point interest rate hike. Estimation from monthly data between January 1968 and December 2006. VAR includes 12 lags. The instrument is the product of the dissent index and the lagged estimated residual in the interest rate equation of the VAR. The first stage F-statistics is 9.53. Point estimates as well as 68 percent and 90 percent confidence intervals are calculated with the code used in Montiel Olea et al. (2020). (95 percent confidence intervals are unbounded in this case)

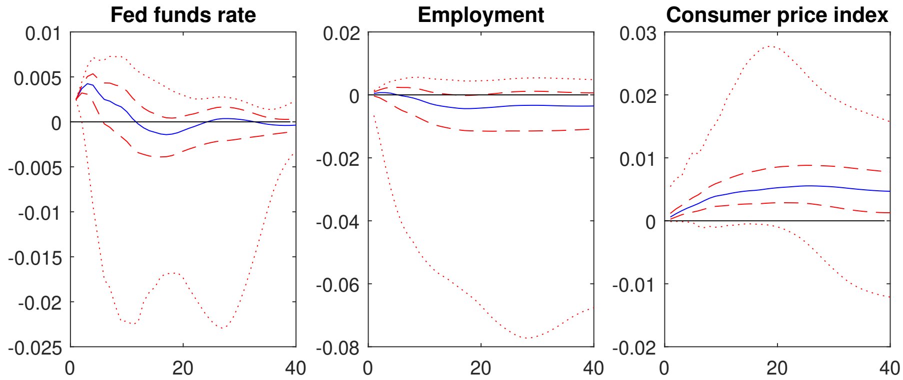

Figure C.4: Effect of an unexpected 25 basis point interest rate hike. Estimation from quarterly data between 1985Q1 and 2006Q4. VAR includes 8 lags. The instrument is the product of the dissent index and the lagged estimated residual in the interest rate equation of the VAR. The first stage F-statistics is 6.73. Point estimates as well as 68 percent and 95 percent confidence intervals are calculated with the code used in Montiel Olea et al. (2020).

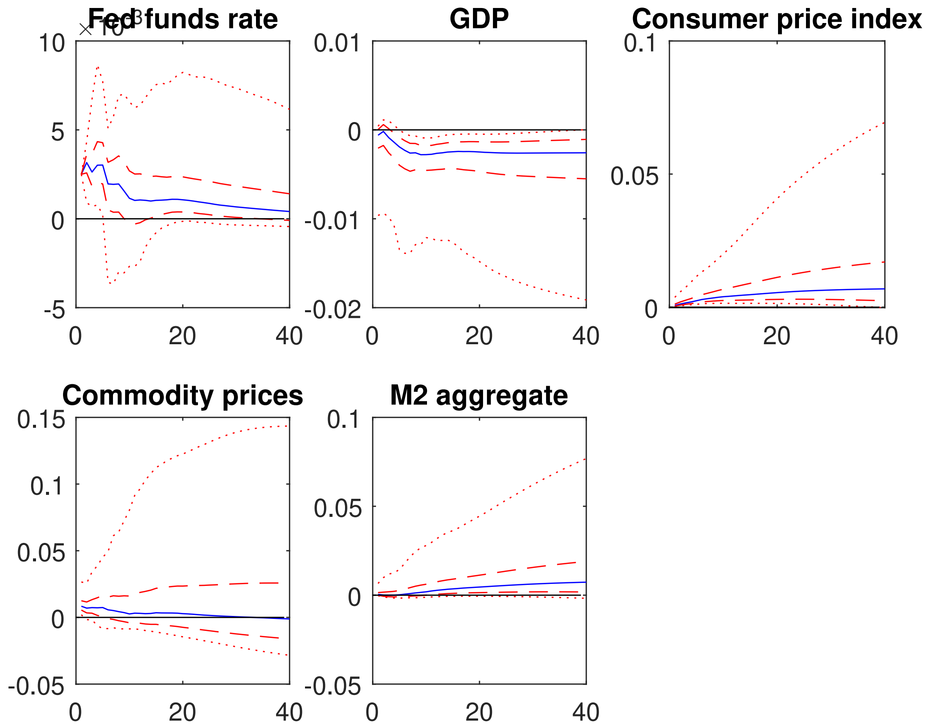

Figure C.5: Effect of an unexpected 25 basis point interest rate hike. Estimation from quarterly data between 1985Q1 and 2006Q4. VAR includes 5 lags. The instrument is the product of the dissent index and the fed funds rate's change in the previous quarter. The first stage F-statistics is 7.31. Point estimates as well as 68 percent and 90 percent confidence intervals are calculated with the code used in Montiel Olea et al. (2020). (95 percent confidence intervals are unbounded in this case)

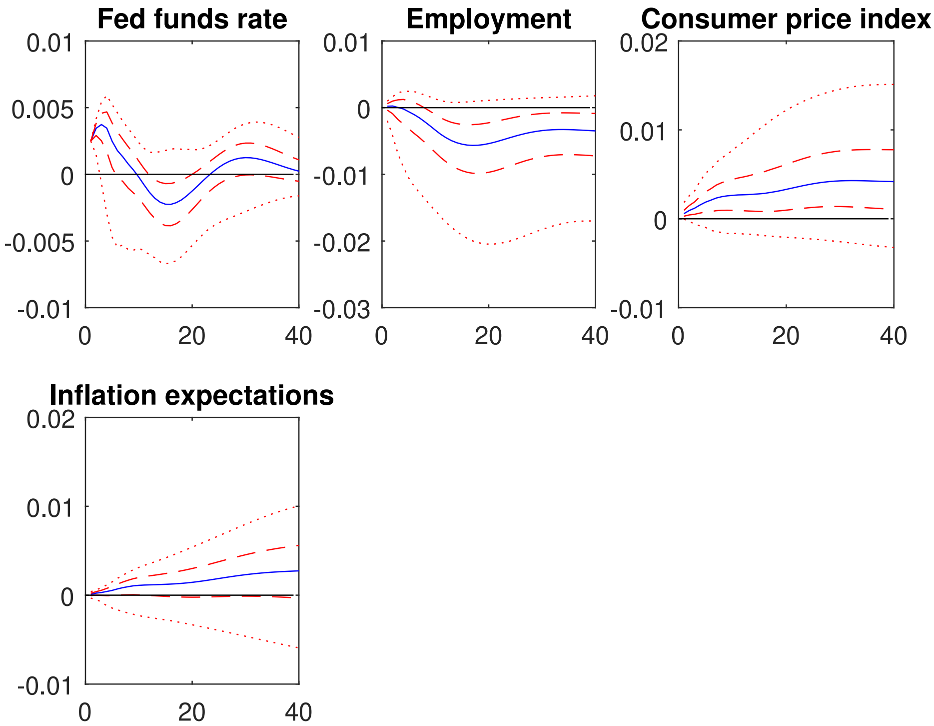

Figure C.6: Effect of an unexpected 25 basis point interest rate hike. Estimation from quarterly data between 1985Q1 and 2006Q4. VAR includes 6 lags. The instrument is the product of the dissent index and the lagged estimated residual in the interest rate equation of the VAR. The first stage F-statistics is 7.79. Point estimates as well as 68 percent and 95 percent confidence intervals are calculated with the code used in Montiel Olea et al. (2020).

Bibliography

Angrist, J., Jordà, Ò. & Kuersteiner, G. (2018). Semiparametric Estimates of Monetary Policy Effects: String Theory Revisited. *Journal of Business & Economic Statistics*, 36:3, 371-387.

Ball, L. & Mankiw, N. G. (1994). Asymmetric price adjustment and economic fluctuations, *Economic Journal*, 104: 247-261.

Barnichon, R. & Matthes, C. (2018). Functional Approximation of Impulse Responses, *Journal of Monetary Economics*, Volume 99: 41-55.

Barth, M. J. & Ramey, V. (2002). The Cost Channel of Monetary Transmission, p. 199-256 in , *NBER Macroeconomics Annual* 2001, Volume 16, National Bureau of Economic Research, Inc,

Bernanke, B. & Gertler, M. (1989). Agency costs, net worth, and business cycles, *American Economic Review*, 79: 14-31.

Bernanke, B. & Mihov, I. (1998). Measuring Monetary Policy. *The Quarterly Journal of Economics*, Volume 113, Issue 3: 869–902.

Boehm, C. & Pandalai-Nayar, N. (2020). Convex Supply Curves, *NBER Working Papers* 26829, National Bureau of Economic Research, Inc.

Caldara, D., & Herbst, E. (2019). Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs. *American Economic Journal: Macroeconomics*, 11 (1): 157-92.

Castelnuovo, E. & Surico, P. (2010). Monetary Policy, Inflation Expectations and The Price Puzzle. *The Economic Journal*, 120: 1262-1283.

Clark, P., Laxton, D., & Rose, D. (1996). Asymmetry in the U.S. Output-Inflation Nexus. *IMF Staff Papers*, International Monetary Fund, 43(1): 216-251.

Cogley, T., & Sargent, T. (2005). Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics*, Volume 8, Issue 2: 262-302.

Coibion, O. (2012). Are the Effects of Monetary Policy Shocks Big or Small? *American Economic Journal: Macroeconomics*, 4 (2): 1-32.

Dolado, J., María-Dolores, R. & Naveira, M. (2005). Are monetary-policy reaction functions asymmetric?: The role of nonlinearity in the Phillips curve, *European Economic Review*, Volume 49, Issue 2: 485-503.

Fan, J. & Gijbels, I. (1996). Local Polynomial Modelling and Its Applications. *Monographs on Statistics and Applied Probability* 66, Chapman and Hall

Fernández-Villaverde, J., Gordon, G., Guerrón-Quintana, P. & Rubio-Ramírez, J. F. (2015). Nonlinear adventures at the zero lower bound, *Journal of Economic Dynamics and Control*, Elsevier, vol. 57(C): 182-204.

Garín, J., Lester, R. & Sims, E. (2018). Raise Rates to Raise Inflation? Neo-Fisherianism in the New Keynesian Model. *Journal of Money, Credit and Banking*, 50: 243-259.

Gerlach-Kristen, P. (2004). Is the MPC's Voting Record Informative about Future UK Monetary Policy?. *Scandinavian Journal of Economics*, 106: 299-313.

Gerlach-Kristen, P. (2008). The Role of the Chairman in Setting Monetary Policy: Individualistic vs. Autocratically Collegial MPCs, *International Journal of Central Banking*, vol. 4(3): 119-143.

Gertler, M. & Karádi, P. (2015). Monetary Policy Surprises, Credit Costs, and Economic Activity. *American Economic Journal: Macroeconomics*, 7 (1): 44-76.

Hirano, K., & Imbens, G. (2004). The Propensity Score with Continuous Treatments. in Gelman, A.& Meng, X-L. (ed) *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*. John Wiley & Sons, Ltd

Imai, K. & van Dyk, D. A. (2004). Causal Inference With General Treatment Regimes, *Journal of the American Statistical Association*, 99:467, 854-866.

Jarociński, M., & Karádi, P. (2020). Deconstructing Monetary Policy Surprises — The Role of Information Shocks. *American Economic Journal: Macroeconomics*, 12 (2): 1-43.

Jordà, Ò. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review*, 95 (1): 161-182.

Mertens, K. & O. Ravn, M. (2013). The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States, *American Economic Review*, 103(4): 1212–1247.

Montiel Olea, J. L., Stock, J. H. & Watson, M. W. (2020), Inference in Structural Vector Autoregressions identified with an external instrument, *Journal of Econometrics*, 2020, ISSN 0304-4076

Primiceri, G. (2005). Time Varying Structural Vector Autoregressions and Monetary Policy, *The Review of Economic Studies*, Volume 72, Issue 3: 821–852.

Rahman, S. & Serletis, A. (2010). The asymmetric effects of oil price and monetary policy shocks: A nonlinear VAR approach. *Energy Economics*, Volume 32, Issue 6: 1460-1466.

Riboni, A. & Ruge-Murcia, F. (2010). Monetary Policy by Committee: Consensus, Chairman Dominance, or Simple Majority?, *The Quarterly Journal of Economics*, 125, issue 1: 363-416.

Riboni, A. & Ruge-Murcia, F. (2014). Dissent in monetary policy decisions, *Journal of Monetary Economics*, Volume 66, 2014: 137-154.

Rubio-Ramirez, J., Waggoner, D. & Zha, T. (2010). Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference. *Review of Economic Studies*, 77, issue 2: 665-696.

Rusnak, M., Havranek, T. & Horvath, R. (2013). How to Solve the Price Puzzle? A Meta-Analysis. *Journal of Money, Credit and Banking*, 45: 37-70.

Santoro, E., Petrella, I., Pfajfar, D. & Gaffeo, E. (2014). Loss aversion and the asymmetric transmission of monetary policy, *Journal of Monetary Economics*, Volume 68: 19-36.

Schaling, E. (2004). The Nonlinear Phillips Curve and Inflation Forecast Targeting: Symmetric versus Asymmetric Monetary Policy Rules. *Journal of Money, Credit and Banking*, 36(3): 361-386.

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

Sims, C. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review*, Volume 36, Issue 5: 975-1000.

Sims, C. A. (2011). Stepping on a rake: The role of fiscal policy in the inflation of the 1970s, *European Economic Review*, Elsevier, vol. 55(1): 48-56.

Stock, J. H., & Watson, M. W. (2012). Disentangling the Channels of the 2007-2009 Recession, *NBER Working Paper Series* 18094, National Bureau of Economic Research, Inc.

Stock, J. H., Wright, J. H. & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20 (4): 518–29.

Taylor, J. B. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, Volume 39: 195-214.

Tenreyro, S. & Thwaites, G. (2016). Pushing on a String: US Monetary Policy Is Less Powerful in Recessions. *American Economic Journal: Macroeconomics*, 8 (4): 43-74.

Thornton, D. L. & Wheelock, D. C. (2014). Making sense of dissents: a history of FOMC dissents, *Review, Federal Reserve Bank of St. Louis*, vol. 96(3): 213-227.

Uhlig, H. (2005). What are the effects of monetary policy on output? Results from an agnostic identification procedure. *Journal of Monetary Economics*, Volume 52, Issue 2: 381-419.

Uribe, M. (2018). The Neo-Fischer effect: econometric evidence from empirical and optimizing models. *NBER Working Paper Series* 25089, National Bureau of Economic Research, Inc.

Weise, C. (1999). The Asymmetric Effects of Monetary Policy: A Nonlinear Vector Autoregression Approach. *Journal of Money, Credit and Banking*, 31(1): 85-108.