Identifying Structural VARs from Sparse Narrative Instruments

Dynamic Effects of U.S. Macroprudential Policies

Katarzyna Budnik*  Gerhard Rünstler†
European Central Bank  European Central Bank

April 2022

Abstract

We study the identification of policy shocks in Bayesian proxy VARs for the case that the instrument consists of sparse qualitative observations indicating the signs of certain shocks. We propose two identification schemes, i.e. linear discriminant analysis and a non-parametric sign concordance criterion. Monte Carlo simulations suggest that these provide more accurate confidence bounds than standard proxy VARs. Our application to U.S. macroprudential policies finds persistent effects of capital requirements and mortgage underwriting standards on credit volumes and house prices together with moderate effects on GDP and inflation.

Keywords: proxy VAR, narrative VAR, capital requirements, borrower-based measures

JEL Codes: C32, E58

* katarzyna.budnik@ecb.europa.eu
† gerhard.ruenstler@ecb.europa.eu. European Central Bank, Sonnemannstrasse 20, D-60314 Frankfurt am Main, Germany. Tel +49 151 61 229 676. Corresponding author. We are grateful to Morten Ravn, Juan Rubio-Ramirez, and the participants of the 2019 ICMAIF and EC2 conferences, the 2019 macro workshop in Gent, the 2020 SVAR workshop at the King’s College London, the 2021 AEA poster session, and various ESCB seminars for comments on earlier versions of the paper. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the European Central Bank.
1 Introduction

Proxy variables and narrative data have been widely used in recent years for identifying policy innovations in vector autoregressive models (VARs). The common principle is to employ them as external information when extracting the innovations from the VAR forecast errors. Optimally, a quantitative proxy variable for the innovation is at hand. For instance, studies have identified monetary policy shocks from high-frequency financial market data indicating the news content of monetary policy communication (Gertler and Karadi, 2016). In areas, where policy actions are difficult to quantify, instruments yet often relate only to a small number of events or are of a qualitative nature. These limitations have been taken up by a narrative approach arguing that qualitative information on just a few events achieves robust identification. Antolin-Diaz and Rubio-Ramirez (2018) show that restrictions on the sign and size of a single policy innovation in October 1979 add important information to identifying U.S. monetary policy shocks in combination with sign restrictions on impulse responses.¹

In this paper, we study the intermediate case of a sparse binary instrument with a possibly larger number of events. We aim at adapting existing Bayesian methods to this case and thereby to bridge a gap between proxy and narrative Bayesian VARs. Following Antolin-Diaz and Rubio-Ramirez (2018) we assume that the econometrician knows no more than the signs of innovations for a limited number of events. However, we also allow for errors in the econometrician’s beliefs, a case that is likely to occur once the number of events increases. We therefore impose narrative restrictions on the expected values of those innovations. We explore two estimation methods for this case. First, we show that the Bayesian proxy VAR can be adapted to a binary instrument from using a discriminant (DC) regression at the identification step. Second, we augment the narrative sign restrictions of Antolin-Diaz and Rubio-Ramirez (2018) with a prior on the degree of sign concordance (SC) to cope with imperfect sign concordance between innovations and the instrument.

¹See also Caldara and Herbst (2019) and Jarocinski and Karadi (2019) for monetary policy and Mertens and Ravn (2013) and Mertens and Montiel-Olea (2018) for fiscal policy applications of proxy VARs. Ludvigson, Ma, and Ng (2017) and Zeev (2018) present applications of the narrative approach.
These modifications may be useful when policy interventions are infrequent and difficult to quantify, as is the case with various types of fiscal and structural policies. While binary instruments may appear less than ideal, our Monte Carlo simulations show that actual efficiency losses are small when moving from quantitative to binary information. At the same time, both DC and SC restrictions provide proper inference with binary indicators. The SC prior is somewhat less efficient than the DC regression, but useful in combination with the latter for safeguarding against errors in the instrument. Among alternative methods, we find that frequentist proxy VARs based on sparse binary indicators are as efficient as the DC regression, but overestimate uncertainty bands, whereas local projections are clearly less efficient.

In our application, we study the effects of macroprudential policies in the post-war U.S. building on a narrative dataset of Elliot, Feldberg, and Lehnert (2013). Just like our methods, a large part of the macro-econometric literature on the impact of macroprudential policies relies on binary narrative indicators, as the high diversity of policy actions impedes the construction of quantitative measures. Reflecting the limited information, studies typically focus on estimating the short-run responses of credit volumes and house prices from panel regressions (see Galati and Moessner, 2017). Our results add to a sporadic literature on the broader macroeconomic dynamics triggered by policy interventions. We focus on capital requirements and mortgage underwriting standards. For both types of policies we find large and highly persistent effects on credit volumes and house prices, together with moderate effects on GDP, inflation, and corporate bond spreads. Long transmission lags suggest that panel regressions focusing on the short run may understate the effects of macroprudential policies. Our findings also relate to the literature on credit supply shocks (Gilchrist and Zakrajsek, 2012; Mian, Sufi and Verner, 2017) and government mortgage purchases (Fieldhouse, Mertens, and Ravn, 2017), and underpin the role of collateral constraints in generating the highly persistent leverage cycles found by Claessens, Kose, and Terrones (2015) and Rünstler and Vlekke (2018).

The remainder of the paper is organised as follows. Section 2 introduces the identification schemes for binary instruments. Sections 3 and 4 present the Monte Carlo simulation exercise and the application to U.S. macroprudential policies, respectively. Section 5 concludes.
2 A Bayesian VAR for Sparse Binary Instruments

Consider the reduced-form VAR for $n \times 1$ vector $x_t$ over periods $t = 1, \ldots, T$,

$$x_t = c + \sum_{s=1}^{r} B_s x_{t-s} + u_t,$$

(1)

where $c$ is a constant term and the VAR forecast errors $u_t$ are assumed to be independent over time with mean $\mathbb{E}u_t = 0$ and covariance matrix $\mathbb{E}u_t u_t^T = \Sigma$. The key assumption of the model is that the VAR forecast errors have a structural representation that isolates a certain innovation $\theta_t$ of interest as the first element of a vector of orthogonal innovations $\epsilon_t$, such that

$$A_0 u_t = \epsilon_t = \begin{pmatrix} \theta_t \\ \epsilon_t^+ \end{pmatrix},$$

(2)

with $\mathbb{E}\epsilon_t \epsilon_t^T = I_n$. Denote with $\alpha^T$ the first row of matrix $A_0$, implying $\alpha^T u_t = \theta_t$. Once $\alpha$ is known, the dynamic response of $x_t$ to innovation $\theta_t$ can be obtained. Proxy and narrative VARs aim at identifying $\alpha$ from outside information about realisations of $\theta_t$. This information may take different forms requiring different statistical methods for estimating $\alpha$. In particular, proxy VARs as introduced by Mertens and Ravn (2012) rely on a quantitative instrument for innovations $\theta_t$, whereas narrative sign restrictions as proposed by Antolin-Diaz and Rubio-Ramirez (2018) impose restrictions on the sign of $\theta_t$ for specific events.

In this paper we consider a weaker version of narrative sign restrictions. Instead of imposing them directly on innovations, we apply them to the expected values of the latter. We thus assume that the econometrician knows about the presence of a mean shift in innovations $\theta_t$ for a set of $m < T$ events, but is ignorant about its size. This defines an instrument $z_t$ that takes values of $z_t = \text{sign}(\mathbb{E} \theta_t)$ for events and $z_t = 0$ otherwise, together with the conditions

$$\mathbb{E}(\theta_t | z_t \neq 0) = \gamma z_t$$

$$\mathbb{E}(\epsilon_t^+ | z_t) = 0,$$

(3)
where $\gamma = \mathbb{E}(\theta_t z_t | z_t \neq 0)$ is the (unknown) expected absolute value of innovations $\theta_t$ for non-zero $z_t$. Given $\mathbb{E}\epsilon_t = 0$ the conditions define a mean shift in $\theta_t$ for $z_t \neq 0$.

Applying the narrative restrictions to expected values makes them suitable for a larger number of events, as it allows for errors in the econometrician’s beliefs, which are likely to occur in this case. As a result, instrument $z_t$ acts as a treatment effect, identifying vector $\alpha$ from a mean shift that is present in $\theta_t$, but not in the remaining innovations $\epsilon_t^\dagger$. At the same, narrative sign restrictions on innovations would still hold with a certain probability, depending on distributional assumptions and the value of $\gamma$.

Stock and Watson (2018) discuss the requirements for the identifiability of $\alpha$ in the general context of proxy VARs, of which the above model is a special case. First, the relevance condition $\gamma > 0$ ensures that instrument $z_t$ picks up events that generate relevant innovations. Second, the exogeneity condition $\mathbb{E}(\epsilon_t^\dagger | z_t) = 0$ requires that events included in $z_t$ are independent of other contemporaneous shocks in the system. Third, while the exogeneity condition is considerably weaker than the requirement of lag exogeneity in regression-based approaches, the latter is replaced by the invertibility condition that innovations $\theta_t$ are fully spanned by the VAR forecast errors. One important reason for non-invertibility is that relevant variables are missing. In practice, the validity of the invertibility condition is therefore determined by the selection of variables $x_t$ included in the VAR.

Estimation amounts to a classification problem. The task is to find linear combinations $\alpha^T u_t$ that segregate the VAR forecast errors in periods with non-zero $z_t$ from those in the remaining periods to achieve mean shift $\gamma z_t$. We explore two ways of estimating parameter vector $\alpha$. First, the discriminant ($DC$) regression estimates $\alpha$ under the assumption that the VAR forecast errors are normally distributed conditional on $u_t$. Second, we augment the narrative sign restrictions of Antolin-Diaz and Rubio-Ramirez (2018) with a prior on the degree of sign concordance ($SC$) to allow for imperfect sign concordance. The $DC$ regression achieves point identification and parallels the proxy VAR approach, while the $SC$ restriction is a non-parametric method based on set identification. We denote the data set with $X = (x_1^T, \ldots, x_T^T)^T$ and the coefficients of the reduced form VAR (1) with $B_\ell = (c, B_1, \ldots, B_p)$. 
2.1 Fisher Discriminant Regression

One way to estimate \( \alpha \) is the discriminant (DC) regression due to Fisher (1931), which is reviewed in Maddala (2013:18ff). The DC regression is designed to predict binary observations \( z_t \) from a set of explanatory variables, which are normally distributed conditional on \( z_t \). We therefore impose the distributional assumptions

\[
\theta_t \mid z_t \sim N(\gamma z_t, \sigma_z),
\]

\[
\epsilon_t^+ \mid z_t \sim N(0, I_{n-1}).
\]

Further, since the cases of \( z_t = +1 \) and \( z_t = -1 \) are symmetric, our classification problem can be transformed into a purely binary one by abstracting from the sign of \( z_t \). Let \( \delta_t = -1 \) if \( z_t = -1 \) and \( \delta_t = 1 \) otherwise and define \( z_t^* = \delta_t z_t - m/T \) and \( u_t^* = \delta_t u_t \). The transformation includes a mean adjustment to ensure a zero sample mean of \( z_t^* \). Hence, \( z_t^* \) takes the values \( z_t^* = 1 - m/T \) for \( z_t \neq 0 \) and \( z_t^* = -m/T \) otherwise. Under the conditions

\[
u_t^* \sim N(\mu_0, \Sigma_z) \quad \text{for} \quad z_t^* = -m/T
\]

\[
u_t^* \sim N(\mu_1, \Sigma_z) \quad \text{for} \quad z_t^* = 1 - m/T
\]

the DC regression

\[
z_t^* = a^T u_t^* + \xi_t
\]

provides an efficient estimate of \( \alpha \) to predict \( z_t^* \) from \( u_t^* \) based on the rule \( \hat{z}_t^* = 1 \) if \( \alpha^T u_t^* > 0 \) and \( \hat{z}_t^* = 0 \) otherwise. Despite the non-standard distribution of \( \xi_t \), the OLS estimate of \( \alpha \) is subject to standard inference as the regression compares the means of two conditional normal distributions, maximising the squared mean difference between the two groups over the variance within groups (see Maddala, 2013:18ff). Assuming an uninformative prior for \( \alpha \) and a Jeffrey prior for \( \sigma_\xi^2 \) gives \( a \mid B_+, \sigma_\xi^2, X, Z \sim N(\hat{a}, \sigma_\xi^2 S_u^{-1}) \) and \( \sigma_\xi^2 \mid B_+, X, Z \sim IG(\hat{\sigma}_\xi, T - n - 1) \), where \( \hat{a} \) and \( \hat{\sigma}_\xi^2 \) are the OLS estimates of equation (6) and \( S_u = \Sigma_{t=1}^T u_t u_t^T \).

The DC regression is a special case of discriminant analysis and emerges as the efficient
solution to our classification problem under an intuitive loss function.\footnote{Discriminant analysis refers to a general theory of classifying categorial observations from quantitative variables based on certain loss functions, see Supplement B.1 for a brief review.} Specifically, under assumptions (4), the DC regression provides the maximum likelihood estimate of vector $a$ under the loss function $mC_1 = (T - m)C_0$, where $C_i$ is the cost of misclassifying an observation with $\delta_t z_t = i$. The cost of misclassification is inversely proportional to the number of observations in each category imposing a high cost of misclassifying non-zero $z_t$ under small $m$, which we regard as a desired feature. While logistic regression is a possible alternative, it is less efficient than the DC regression with normally distributed explanatory variables (Efron, 1975). It is also more difficult to implement, as it may suffer from convergence issues due to complete separation of the likelihood function, a case that is likely to occur with small $m$ (Allison, 2008).\footnote{We did faced this issue when experimenting with Bayesian logistic regressions in our application.}

At the same time, the DC regression has been found to be fairly robust against moderate deviations from normality (Pohar, Blas, and Turk, 2003; Maddala, 2013:18ff).

### 2.2 Sign Concordance

A useful statistics for the relevance of instrument $z_t$ is the sign concordance (SC) statistics $\phi$, defined as the share of instances for which the sign of $\theta_t$ coincides with $z_t$,

$$\varphi(B_+, \alpha, X, Z) = m^{-1} \sum_{z_t \neq 0} I(\theta_t z_t > 0),$$

where $I()$ denotes the indicator function. Given the independence of $\theta_t$ over time, the number of correct signs follows a binomial distribution,

$$p(m\varphi|\alpha, \lambda, B_+, X, Z) = f_z(m\varphi; m, \lambda),$$

where $\lambda$ is the unknown probability of the correct classification of a single event.

In combination with an appropriate prior on $\lambda$ that supports acceptance of high values of $\varphi$, the SC statistics can be used as a non-parametric alternative to the DC regression for estimating $\alpha$ based on set identification. The principle is to use the prior as a weighting
scheme for accepting uninformative draws of \( \alpha \) dependent on the value of \( \varphi \). A flexible choice is a beta-distribution \( \lambda \sim \beta(p, q) \) over support \([\Lambda, 1]\) with \( \Lambda > 0.5 \). Figure 1 shows examples of the resulting beta-binomial prior \( f(\varphi; m, \Lambda, p, q) \) for \( \varphi \) for different values of \( m, \Lambda, \) and \((p, 1)\). The benchmark case of a uniform distribution, \( p = q = 1 \), creates steadily increasing acceptance weights with a smooth threshold at \( \Lambda \). Weights increase more sharply with \( p > 1 \).

The above principle may be applied to any acceptance criterion that takes a binary form for individual events. One example is the narrative restrictions based on relative magnitudes in the historical decomposition of the VAR as used by Antolin-Diaz and Rubio-Ramirez (2018). Moreover, one may generate more complex priors that allow for different acceptance probabilities across individual events by means of the Poisson binomial distribution. The computational complexity increases yet rapidly with \( m \). Chen and Liu (1997) discuss various efficient methods to obtain the density of this distribution.

### 2.3 Sampling

The distributional assumptions (4) underlying the DC regression imply a non-standard unconditional distribution of the VAR forecast errors, as it emerges as a mixture of two normal distributions with different means. However, normality can be established conditional on mean shift \( \gamma \) by adding the contemporaneous impact of the latter as a deterministic term to the VAR. This allows for estimating the model via a Gibbs sampler iterating between the reduced form VAR and the DC regression. We assume a standard Normal-Wishart prior for parameters \( B^+ \) and \( \Sigma_z \) and uniformative priors for \( \alpha \) and \( \gamma \). Sampling proceeds as follows:

1. Draw from the posterior of \( a|B_+ \) as described in section 2.1. Rescale \( \alpha = (a^T S_u a)^{-1} a \) to achieve unit variance of \( \theta_t = \alpha^T u_t \).

Construct matrix \( A_0 \) from \( A_0^T = A_s Q \), where \( A_s \) is the Choleski decomposition of \( S_u \) and \( Q = (q_1, \ldots, q_n) \) is an orthonormal matrix, \( QQ^T = I_n \). Set \( q_1 = A_s^{-1} \alpha \) such that the first row of \( A_0 \) contains \( \alpha^T \). The remaining columns \( q_2, \ldots, q_n \) are irrelevant and may be constructed from a Gram-Schmidt orthogonalisation as in Arias et al. (2018).
(2) Draw from the posteriors of $\gamma|B_+, \alpha, \sigma^2_\gamma \sim N(\hat{\gamma}, m^{-1}\sigma^2_\gamma)$ and $\sigma^2_\gamma|B_+ , \alpha \sim IG(\hat{\sigma}^2_\gamma, m - 1)$, where $\hat{\gamma} = m^{-1} \sum_{t=1}^{T} \theta_t z_t$ and $\hat{\sigma}^2_\gamma = m^{-1} \sum_{t=1}^{T} (\theta_t z_t)^2 - \hat{\gamma}^2$.

(3) Obtain the contemporaneous impact $\Delta z_t$ of mean shift $\gamma$ on $x_t$, where

$$\Delta = \mathbb{E}(u_t|z_t = 1) - \mathbb{E}(u_t|z_t = 0) = \left(1 + \frac{m^+ - m^-}{T - m}\right) A_0^{-1} \begin{pmatrix} \gamma \\ 0_{n-1} \end{pmatrix},$$

$m^+$ and $m^-$ are the numbers of positive and negative values of $z_t$, respectively, and $0_{n-1}$ is a zero vector of length $n - 1$ (see Supplement B.2).

(3) Draw from the posterior of $B^+, \Sigma_z|\alpha, \gamma$ based on the VAR

$$x_t - \Delta z_t = c + \sum_{s=1}^{r} B_s x_{t-s} + u_t^\Delta,$$

where $u_t^\Delta \sim N(0, \Sigma_z)$. Obtain the forecast errors $u_t = u_t^\Delta + \Delta z_t$ of the VAR (1).

In case of SC prior we follow Antolin-Diaz and Rubio-Ramirez (2018) in drawing from the posterior of $\alpha$ by rejection sampling. Steps (1) and (2) from above are replaced as follows.

(1’) Obtain an uninformative draw of $\alpha$. Following Arias et al. (2018) we specify $\alpha = A_* q_1$, where $q_1$ is a draw from the Haar measure of orthogonal matrices. This is obtained as $q_1 = v/||v||$ from a random draw of vector $v \sim N(0, I_n)$. Construct the remaining columns $q_2, \ldots, q_n$ from a Gram-Schmidt orthogonalisation as in Arias et al. (2018).\(^5\)

(2’) Draw from the prior of $\lambda$ and accept the draw with probability $f_z(m\varphi; m, \lambda)$.

Note that the SC prior does not require specific distributional assumptions. It is therefore consistent with the assumption of an unconditional normal distribution of the VAR forecast errors as in Antolin-Diaz and Rubio-Ramirez (2018). In this case, the mean adjustment $\Delta$ is ignored and the Gibbs sampler collapses to direct sampling. Alternatively, as discussed in section 2.4 below, the DC regression may be combined with the SC prior. This can be

\(^4\)The adjustment term arises, as $\mathbb{E}u_t = 0$ and $\mathbb{E}(u_t|z_t \neq 0) > 0$ imply a non-zero expectation $\mathbb{E}(u_t|z_t = 0)$.

\(^5\)The Haar measure has been subject to controversy. Giacomini et al. (2021) propose robust priors as an alternative. Inoue and Kilian (2020) argue that concerns about the Haar measure have been overstated.
implemented from drawing $\alpha$ as from the $DC$ regression and adding the rejection sampling step (2') after step (2). Finally, $DC$ and $SC$ restrictions may be integrated in the framework of Arias, Rubio-Ramirez, and Waggoner (2018) to combine them with sign and zero restrictions on impulse responses. Further technical details are discussed in Supplement B.2.

### 2.4 Comparison with Existing Approaches

The $DC$ regression may be viewed as an adaption of a standard Bayesian proxy VAR to the case of a binary instrument. Proxy VARs are defined as in equations (1) and (2). They assume a quantitative instrument and achieve identification from the moment condition $E\theta_t z_t > 0$, together with $n - 1$ orthogonality conditions $E\epsilon_t^+ z_t = 0$. An estimate of $\alpha$ is obtained from the proxy regression $z_t = a^T u_t + \xi_t$ (see e.g. Caldara and Herbst, 2019). Conditions (3) are a special case of the proxy VAR assumptions, while $DC$ and proxy regressions are similar to each other. The key difference is the distributional assumption on the residual $\xi_t$ of the regressions. Bayesian proxy VARs so far have maintained the assumption that $\xi_t$ follows a standard normal distribution (Caldara and Herbst, 2019; Giacomini et al. 2021), which clearly does not apply to sparse instruments.$^6$ The $DC$ regression achieves his adjustment. Applications of frequentist proxy VARs typically use bootstrap techniques for estimating confidence bands to cope with the fact that estimation proceeds in two steps. These bootstraps are, in principle, capable of dealing with a non-standard distribution of $\xi_t$. However, Jentsch and Lunsford (2016) show that standard methods grossly underestimate confidence bands if instruments are sparse. They propose a modified block bootstrap for this case, whereas Montiel Olea, Stock, and Watson (2020) present bootstraps that are also suitable for weak instruments. Note that the application of the $DC$ regression in a frequentist context would require a bootstrap as well.

The $SC$ prior is a generalisation of narrative restrictions to the case of incomplete sign concordance. The identification conditions used in the related studies are as in section 2.1 (Antolin-

$^6$Arias et al. (2021) take a different route implementing the moment conditions as deterministic restrictions.
Diaz and Rubio-Ramirez, 2018; Ludvigson et al., 2017; Zeev, 2018). These papers use very small numbers of events and assume perfect sign concordance. Their restrictions are therefore a special case of the SC prior with $\lambda = 1$. While the case of $\lambda = 1$ can, in principle, be applied to a larger number of events, it may turn out overly tight in practice.

One important difference between DC and SC restrictions is that the former is parametric and achieves point identification by maximising a function of the mean shift $\gamma$, whereas the latter is a non-parametric approach resulting in set identification. While point identification is likely to result in narrower confidence bands, the SC restriction may be more robust against outliers or other violations of distributional assumptions on the VAR forecast errors. Moreover, the SC restriction may be combined with other types of restrictions, such as sign restrictions on the impulse responses to innovation $\theta_t$ as in Antolin-Diaz and Rubio-Ramirez (2018).

In fact, DC and SC restrictions may be combined with each other. In this case the SC prior attains an interpretation as reliability prior, as proposed by Caldara and Herbst (2019) for a proxy VAR with standard distributional assumptions. The prior then acts to give higher weight to draws of $\alpha$ from the DC regression that result in high values of sign concordance. Note that the prior feeds back into draws of the reduced-form VAR coefficients regulating the informativeness of the instrument for the posterior distribution of the latter.

## 3 Sparse Policy Interventions: a Monte Carlo Study

This section presents Monte Carlo simulations to compare DC and SC restrictions with frequentist proxy VARs and local projections. We inspect the bias and uncertainty in the estimates of impulse responses (IRFs) together with the accuracy of uncertainty bounds.

The simulations extend upon the econometric framework of section 2 by assuming that innovations $\theta_t$ include a sparse set of random policy interventions $\zeta_t$, which we use to generate instrument $z_t$. Our setup allows us to study cases where $z_t$ is a contaminated measure of actual policy interventions with either irrelevant events being added to the instrument or rel-
relevant events missing from the latter. Such trade-off is important in application as the scope of interventions to be included in a sparse instrument is not well-defined given the limited ex-ante knowledge about their size: for instance, some interventions may have not been binding or may be largely explained by past developments. The researcher therefore faces a choice between including many interventions, of which some may be of limited relevance, or focusing on a small set of major events. Mertens and Montiel-Olea (2018) advocate the latter strategy.

We use the data generating process

\[
\begin{pmatrix}
    x_{t,1} \\
    x_{t,2}
\end{pmatrix}
= B_1
\begin{pmatrix}
    x_{t-1,1} \\
    x_{t-1,2}
\end{pmatrix}
+ A_0^{-1}
\begin{pmatrix}
    \eta_t \\
    \zeta_t
\end{pmatrix},
\]

where innovations \( \theta_t = \eta_t + \zeta_t \) are the sum of regular innovations \( \eta_t \) and independent sparse policy interventions \( \zeta_t \). We let \( (\eta_t, \epsilon_t^\pm)^T \sim N(0, 10^{-2}I_2) \). Sparse policy interventions \( \zeta_t \) are generated by the policy rule

\[
\begin{align*}
\zeta_t^* &= \omega x_{t-1,2} + \nu_t \\
\zeta_t &= -\mathbb{I}(\zeta_t^* \geq \bar{\zeta})\zeta_t^+.
\end{align*}
\]

Interventions \( \zeta_t \) arise both from exogenous shocks \( \nu_t \) and the dependency of the policy target \( \zeta_t^* \) on the past state of the system. The policy-maker intervenes only once \( \zeta_t^* \) exceeds a certain threshold \( \bar{\zeta} \). All interventions \( \zeta_t \) are negative. Their absolute size \( \zeta_t^+ \) is drawn from a lognormal distribution \( \ln(\zeta_t^+/\bar{\zeta}) \sim N(-\sigma_\zeta^2/2, \sigma_\zeta^2) \) such that \( \mathbb{E}\zeta_t^+ = \bar{\zeta} \) and \( \text{var}(\zeta_t^+) = \exp(\sigma_\zeta^2) - 1 \). We set \( T = 200 \) and calibrate \( \bar{\zeta} \) to achieve a number of interventions of either \( m = 10 \) or \( m = 20 \) and set the dispersion of interventions to \( \sigma_\zeta = 0.005 \) or \( \sigma_\zeta = 0.01 \). We choose matrix \( B_1 \) to generate cyclical fluctuations with a length of 32 quarters in \( x_t \) and matrix \( A_0 \) to achieve a correlation of 0.3 among the VAR residuals together with a large initial response of \( x_{1,t} \) to \( \theta_t \).\(^7\)

We compare seven models. As a benchmark serves a Bayesian proxy VAR that uses the true policy interventions \( \zeta_t \) as instrument \((BV\zeta)\). The remaining models use instrument \( z_t \).

\(^7\)More precisely, we set \( \bar{\zeta} \) to values of 0.0164 and 0.0128 for \( m = 10 \) and \( m = 20 \), respectively, and discard draws that do not deliver the desired number of events. The data generating process is further described in Supplement B.3.
We consider the $DC$ regression, a uniform $SC$ prior for $\lambda$ over interval $[0.85, 1]$, and the combination of the $DC$ regression with the $SC$ prior (model $DSC$). In all cases we use an uninformative prior for the reduced-form VAR. Among frequentist models, we include a proxy VAR ($pV$) with uncertainty bands obtained from the bootstrap of Montiel-Olea et al. (2020), local projections and a version of the recursive VAR proposed by Plagborg-Møller and Wolf (2021). The latter two models include $z_t$ directly as a regressor instead of using it as an instrument and use a standard bootstraps. These models are described in detail in Supplement B.3.\(^8\)

The first four simulations shown in Table 1 ignore contamination issues assuming that the econometrician observes $z_t = \text{sign}(\zeta_t)$. Note that the instrument nevertheless might differ from $\text{sign}(\theta_t)$ in certain periods due to innovations $\eta_t$, resulting in imperfect sign concordance. The table shows statistics on the standardised IRF of $x_{t,1}$ including the root mean squared error (RMSE) and bias of the central estimate, the $[0.1, 0.9]$ interquantile difference of its distribution as a measure of its true uncertainty, and the width of the corresponding estimated uncertainty bands. Moreover, Table 1 shows coverage ratios, defined as share of draws where the true IRF lies within the estimated uncertainty bands. Supplement C.1 plots the IRFs together with true and estimated uncertainty bands.

We find, first, that the efficiency losses from using $z_t$ in place of the true policy intervention $\zeta_t$ as in the Bayesian proxy VAR $BV\zeta$ remain modest, unless the dispersion $\sigma_\zeta$ of policy innovations becomes large. Whereas losses are negligible for a value of $\sigma_\zeta = 0.01$, they increase for a value $\sigma_\zeta = 0.01$ as in simulation (2), because the higher $\sigma_\zeta$ implies a higher share of small policy innovations $\zeta_t$ which are confounded by innovations $\eta_t$. Second, in all simulations, the efficiency of the $SC$ prior falls somewhat short of the $DC$ regression, while the estimated IRF uncertainty bands are largely appropriate for models $DC$, $SC$, and $DSC$. At horizon 0, models $DC$ and $SC$ tend to overestimate the width of the bands, whereas model $DSC$ underestimates them. At horizon 4, however, model $DSC$ uniformly provides more accurate estimates of uncertainty bands than model $DC$ reflecting the information gain from the $SC$

\(^8\)For model $pV$ we build on the replication files of Mertens and Montiel Olea (2018).
prior feeding back into the posterior reduced-form VAR coefficients.

Third, while the central estimates from the frequentist proxy VAR are very similar to those from the $DC$ regression by construction, the bootstrap of Montiel-Olea et al. (2020) clearly overestimates uncertainty bands. We obtained this outcome also for the bootstrap of Jentsch and Lumsford (2016). Local projections and the recursive VAR, which include $z_t$ directly as a regressor to estimate the impact of policy innovations, are clearly less efficient than any of the above models. We note that our simulations ignore invertibility issues, which are an advantage of regression-based methods compared to proxy VARs (Stock and Watson, 2018). These efficiency losses are yet to be considered when comparing the various methods.9

The final two simulations deviate from the specification $z_t = \text{sign}(\zeta_t)$ to study the implications of observation errors in instrument $z_t$. We consider the trade-off between the two possibilities that the econometrician either misses relevant interventions in instrument $z_t$ or mistakenly includes redundant interventions. We take the perspective of an econometrician who faces 20 potential policy events, but is ignorant about their relevance. Simulation (5) assumes that the econometrician mistakenly adds 10 redundant events to $z_t$, which have no correspondence to policy shocks $\zeta_t$: we generate $m = 10$ true events $\zeta_t$ and add another 10 random non-zero observations to $z_t$. Simulation (6) studies the case that the econometrician ignores 10 relevant events: we generate $m = 20$ events $\zeta_t$, but remove 10 of those events from $z_t$.

We find that both types of error result in higher uncertainty of the estimates, but redundant events create definitely larger distortions than missing events. The removal of 10 relevant events in simulation (6) creates only a modest increase in the RMSE compared to simulations (1) and (3). By contrast, adding 10 irrelevant events in simulation (5) almost doubles the RMSE as it generates substantial downward biases in IRF estimates. Interestingly, the $SC$ prior performs as well as the $DC$ regression in case of redundant events. In both cases, yet model $DSC$ turns out best, as it provides the most accurate estimates of uncertainty bands. Overall, the results suggest a conservative approach to constructing sparse instruments, while the combination of the $DC$ regression with the $SC$ prior appears to provide some insurance

---

9See Kilian and Kim (2011) and Li et al. (2021) for related studies.
against both types of observations error.

4 Macroprudential Policy Interventions in the U.S.

We apply our approach to estimating the effects of postwar U.S. policy interventions related to capital requirements and underwriting standards on mortgage credit in the postwar U.S. Capital requirements and borrower-based measures (of which underwriting standards are an important category) represent the most widely used macroprudential policy instruments since the 2008 Global Financial Crisis. The impact assessment of these policies is often impeded by the poor quality of data on policy interventions. First, macroprudential policies have been carved out from supervisory and regulatory policies only after the 2008 Crisis. Accordingly, historical databases differ substantially in their coverage of policy interventions due to ambiguity in the ex-post classification of individual interventions as being of a macroprudential nature (Budnik and Kleibl, 2018). Second, interventions involve diverse instruments and are scattered across time, impeding the construction of quantitative measures of the policy stance. The majority of studies on the macroeconomic effects of macroprudential policies therefore use binary indicators in cross-country panel regressions to assess the effects of interventions on credit volumes and house prices.\footnote{See e.g. Vandenbussche et al. (2015), Cerutti, Claessens and Laeven (2017), and Budnik (2020).} Galati and Moessner (2017) note that these studies thereby remain silent on issues such as the persistence of these effects and transmission lags.

In particular, given that a macroprudential tightening constrains credit, it may have macroeconomic costs in terms of lower output and higher interest rates. Balancing these costs against financial stability benefits is key to the design of optimal policies and the co-ordination of macroprudential with monetary policies (van der Ghote, 2021). The macroeconomic dynamics triggered by policy interventions has yet been addressed only by a few empirical studies. Richter, Schularick, and Shim (2019) and Pogoshyan (2020) use panel local projection methods to estimate the effects of borrower-based measures. Both studies report declines in response of credit and house prices over horizons of four years. Richter et al. (2019) also
study GDP and inflation. They find a decline of GDP after a tightening of loan-to-value ratios for emerging economies, but no effect for advanced economies, while inflation tends to increase in both cases. Kim and Mehrotra (2017, 2018) include the count of policy events as an endogenous variable in a structural panel VAR. They identify macroprudential policy shocks from a Choleski decomposition assuming a zero contemporaneous response of these shocks to all other innovations apart from monetary policy. For a panel of 17 emerging and advanced economies they find small negative effects on both GDP and inflation.

We study the macroeconomic effects of macroprudential policies for the U.S over the period of 1958Q1 to 2016Q4. Our VAR for includes seven series: real GDP ($y_t$), consumer prices ($p_t$), the effective Federal Funds Rate ($r_t$), the spread between the rate of return on BAA corporate bonds and the 10-year Treasury Bond ($r^C_t$), real total credit to the non-financial corporate sector ($c^P_t$) and the household sector ($c^H_t$), and real residential property prices ($h_t$). With the exception of interest rates, the series enter the VAR as quarterly log-differences. The data are taken from the FRED database. For residential property prices we use the Shiller U.S. national home price index. The credit data are from the BIS long credit statistics.\footnote{See https://fred.stlouisfed.org/ and https://www.bis.org/statistics/totcredit.htm.}

4.1 The Narrative Indicators

The major source of our information on capital requirements and mortgage underwriting standards is the database of Elliott, Feldberg, and Lehnert (2013), which contains a wide range of policy interventions addressing macro-financial risks in the U.S. in between 1914 and the early 1990s. We augment the information provided by Elliot et al. (2013) until the end of 2016 based on various sources. In particular, we add interventions related to capital requirements introduced after the Basel Accords and Agreements. For both policies we define instrument $z_t$ such that $z_t = -1$ in case an expansionary measure was set in period $t$, $z_t = 1$ in case of a contractionary measure, and $z_t = 0$ otherwise. This results in 10 events each for capital requirements and underwriting standards. The events are listed in Supplement A.
These policy interventions aimed at controlling macro-financial risks. They may therefore not represent exogenous shocks. However, as argued by Richter et al. (2019), a response of macroprudential authorities to macro-financial shocks within the same quarter is rather unlikely. Instead, authorities would respond at a certain point in time to imbalances that have accumulated over the past. Such delayed response is reinforced by the specificities of the U.S. institutional framework, as the responsibility for policy interventions has been distributed over different agencies, including the U.S. Congress. Policy actions typically required multiple consultations rendering the exact timing of policy actions less predictable (Elliot et al., 2013). Hence, condition $E(\epsilon_t^+ | z_t)$ in equations (3) is very likely satisfied.

We examine lagged dependencies of the indicators on the endogenous variables $x_t$ included in the VAR from ordered probit regressions. Table 2 shows the results from likelihood ratio tests of the joint significance of coefficients related to each series. The regressions indicate lagged dependencies of the indicators on their respective main target variables only at higher lags. When considering up to four lags, we find some predictive power of credit to households for capital requirements and the bond spread house prices for underwriting standards. However, these effects vanish, if only the first two lags of $x_t$ are considered. We remove one event from the capital requirements indicator that is correctly classified by the probit.

4.2 Impulse Responses to Policy Shocks

We turn to estimating the impulse responses (IRFs) to macroprudential policy innovations from the narrative VAR. We consider three models, i.e. the DC regression, the SC prior, and the combination of the two criteria in model DSC, which gives the SC criterion an interpretation as reliability prior. We specify the prior for $\lambda$ as a uniform distribution over support $[0.9, 1]$. We use the seven variables described above, include eight lags and impose a standard Minnesota prior on the reduced form VAR based on a standard Normal-Wishart prior for $B_+$ as described by Karlsson (2013).\footnote{We specify the prior variance of coefficient $B_{s,ij}$ as $\tau_{s,ij} = (\pi_0 * s_i^{(-\pi_3)})^2 s_j$, where $s_i$ is the residual variance of an univariate autoregressions of series $y_{i,t}$. We set overall tightness $\pi_0 = 0.2$, lag decay $\pi_3 = 0.5$, and use a}
Figure 2 shows the sign concordance posteriors from the three models. For underwriting standards, the number of correctly classified events peaks at a value of 9 out of 10 events with little difference between $DC$ and $SC$ restrictions. For capital requirements the $DC$ restriction gives rise to a substantial share of draws with a low sign concordance of $\varphi < 0.5$ resulting in a median value of the $SC$ posterior of below 0.7. In both cases, the combination of the $DC$ regression with the $SC$ prior acts to reduce the weight of draws with low $\varphi$ shifting the $SC$ posterior to the right compared to both the $DC$ regression and $SC$ estimates used in isolation.

The impulse responses (IRFs) to a policy tightening in capital requirements and underwriting standards are shown in Figures 4 and 5. IRFs are standardized to give the response to a unity shock. We show results for nominal residential property prices $p_H^t$. The IRF estimates turn out very similar across the three models. In line with the simulation results from section 3, estimates based on the $SC$ prior tend to deliver somewhat smaller median responses and larger confidence bounds compared to models $DC$ and $DSC$ (see Fig. Supplement C.1).

For both measures, a policy tightening induces a persistent decline in credit, while the corporate bond spread is subject to a small, but significant increase. At the same time, economic activity and inflation decline. The effects of the two types of policy measures differ in two ways. First, the impact of a change in capital requirements is concentrated on corporate credit, while leaving household credit and house prices unaffected. By contrast, a change in underwriting standards affects both credit categories and results in a pronounced decline in house prices. Second, the impact of a change in capital requirements is less persistent. The response of corporate credit peaks after about two years, while the effect on economic activity reverses after four quarters. For underwriting standards, the responses of both series stabilize only after about 4 years and are highly persistent.

Table 3 presents the average effect of a policy intervention on the series included in the VAR as estimated from model $DSC$. For each draw we multiply the IRF with the corresponding estimate of $\gamma$, which is obtained as the average of sign-adjusted innovations, $\hat{\gamma} = \sum_{z_t \neq 0} \theta_t z_t$.
We find a larger impact of underwriting standards compared to capital requirements. Interventions on underwriting standards on average resulted in declines of corporate and household credit of 0.8% and 1.2%, respectively, after 32 quarters, while house prices dropped by close to 2.2%. Interventions on capital requirement induced a decline in corporate credit of 0.7% but had little effect on household credit and house prices. In both cases, GDP declined by about 0.3% after a year, while the corporate bond spread increased by close to 10 basis points.

Overall, our estimates on the short-term response of credit and house prices are in line with the literature based on cross-country panel regressions. In a meta-analysis of studies based on cross-country panel regressions, Gadea-Rivas, Bräuer and Perez-Quiros (2019) find an average response of credit volumes of about 0.5% in advanced economies after a year. Pogosyan (2020) reports similar outcomes for credit and house prices in the euro area based on local projections. Medium-term effects have so far only been addressed by Kim and Mehrotra (2017, 2018) for emerging economies. Our estimates for the U.S. indicate long transmission lags in the responses to borrower-based measures and moderate, but significant declines in economic activity. Moreover, the weak response of the housing sector to a shift in capital requirements suggests a shift towards a credit portfolio subject to lower risk weights in the banks’ response to a change in capital requirements.

4.3 Robustness

Our initial concern is the validity of the invertibility condition discussed in section 2. In the main estimates, the instrument is set to non-zero values in those periods when the policy interventions entered into force. However, it cannot be ruled out that the effects of these interventions are only partly spanned by the VAR forecast errors in these periods. As discussed by Stock and Watson (2018), any effect of policy interventions not spanned by the forecast errors at implementation would necessarily materialise in forecast errors in subsequent periods.\footnote{The impact of an event $\Theta_t$ can be expressed as $\mathbb{P}(x_{t+h} - x_t|\Theta_t) = \sum_{s=0}^{h} C_s \mathbb{P}(u_{t+h-s}|\Theta_t)$, where $\mathbb{P}$ is the linear projections operator and $x_t = \sum_{s=0}^{\infty} C_s u_{t-s}$ is the moving average representation of the VAR. The invertibility condition $\mathbb{P}(\Theta_t|u_t) = \theta_t$ implies $\mathbb{P}(u_{t+h-i}|\theta_t) = 0$ for $i \neq h$.}
With a binary instrument, the validity of the invertibility condition can therefore be explored by projecting the instrument on the VAR forecast errors subsequent to the implementation. Table 4 reports the outcome of a related exercise based on model DC, where we set the instrument to non-zero values simultaneously for lags 1 to 4 after the adoption date. We thereby integrate over potential effects at individual lags. We find minor further negative effects on household credit and house prices for both policy measures, but none of these would approach significance or alter our conclusions. We also experimented with setting the instrument to non-zero values at individual lags, using the SC prior as a selection criterion to attach larger weight to lags that generate further effects. This gave similar results. The IRFs are shown in Figure Supplement C.4.

The results of further robustness checks are shown in Figures Supplement C.5 and C.6. We consider estimates based on an uninformative prior for the VAR with 4 lags and add a banking deregulation index to the baseline VAR in order to control for the deregulation of the U.S. banking sector in the 1980s (Mian, Sufi and Verner 2017). Estimates from frequentist proxy and recursive VARs are show some differences compared to model DSC. In particular, the proxy VAR finds an increase in household credit and house prices after a tightening of capital requirements, while the recursive VAR finds a decline. Further, the recursive VAR does not detect an effect of underwriting standards on house prices.

5 Conclusions

This paper discussed identification based on sparse binary instruments in the context of Bayesian VARs. It proposed two estimation methods inspired by proxy and narrative VARs, respectively. These methods should be useful for assessing the effects of policies in areas where interventions are infrequent and difficult to quantify. Our Monte Carlo simulations indicate that efficiency losses from using binary instruments remain contained, while Bayesian methods provide more accurate inference than frequentist proxy VARs and are clearly more efficient than regression-based approaches. A conservative approach to constructing sparse
instruments based on a small set of relevant events is likely to produce sharper results, while the $SC$ reliability prior appears to provide some insurance against errors in the instrument.

Our application to the effects of macroprudential policies in the postwar U.S. indicates long transmission lags in the response of credit and house prices in particular to borrower-based measures. Studies based on cross-country panel regressions that inspect at short horizons may therefore understate the overall effects of policy measures. Moreover, we found moderate, but significant declines in economic activity and increases in corporate bond spreads after a policy tightening indicating a need for coordinating macroprudential with monetary policy. Our findings are also informative about the impact of shifts in credit supply and household collateral constraints in general and underpin the high persistence of leverage cycles documented, for instance, by Claessens et al. (2012) and Rünstler and Vlekke (2018). Similarly, Fieldhouse et al. (2018) have stressed that the easing of borrowing constraints due to financial innovation has materialized in house prices only with long lags over the last housing cycle.
References


<table>
<thead>
<tr>
<th></th>
<th>m</th>
<th>100*σζ</th>
<th>ω</th>
<th>BVζ</th>
<th>DC</th>
<th>DSC</th>
<th>SC</th>
<th>pV</th>
<th>LP</th>
<th>rV</th>
<th>DC</th>
<th>DSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Baseline</td>
<td>10</td>
<td>0.5</td>
<td>0.0</td>
<td>RMSE</td>
<td>.13</td>
<td>.15</td>
<td>.16</td>
<td>.20</td>
<td>.15</td>
<td>.31</td>
<td>.29</td>
<td>.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bias</td>
<td>-.04</td>
<td>-.06</td>
<td>-.07</td>
<td>-.13</td>
<td>-.05</td>
<td>-.02</td>
<td>-.05</td>
<td>-.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IQD</td>
<td>.42</td>
<td>.46</td>
<td>.52</td>
<td>.62</td>
<td>.45</td>
<td>1.03</td>
<td>.94</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UB</td>
<td>.46</td>
<td>.53</td>
<td>.40</td>
<td>.73</td>
<td>.72</td>
<td>1.12</td>
<td>.93</td>
<td>.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Coverage</td>
<td>.81</td>
<td>.83</td>
<td>.71</td>
<td>.89</td>
<td>.88</td>
<td>.82</td>
<td>.78</td>
<td>.74</td>
</tr>
<tr>
<td>(2) Baseline</td>
<td>10</td>
<td>1.0</td>
<td>0.0</td>
<td>RMSE</td>
<td>.15</td>
<td>.18</td>
<td>.20</td>
<td>.25</td>
<td>.18</td>
<td>.41</td>
<td>.37</td>
<td>.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bias</td>
<td>-.01</td>
<td>-.06</td>
<td>-.07</td>
<td>-.17</td>
<td>-.04</td>
<td>-.00</td>
<td>-.03</td>
<td>-.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IQD</td>
<td>.42</td>
<td>.56</td>
<td>.62</td>
<td>.82</td>
<td>.53</td>
<td>1.33</td>
<td>1.21</td>
<td>.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UB</td>
<td>.42</td>
<td>.60</td>
<td>.44</td>
<td>.82</td>
<td>.90</td>
<td>1.42</td>
<td>1.14</td>
<td>.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Coverage</td>
<td>.73</td>
<td>.80</td>
<td>.66</td>
<td>.86</td>
<td>.90</td>
<td>.80</td>
<td>.77</td>
<td>.72</td>
</tr>
<tr>
<td>(3) Baseline</td>
<td>20</td>
<td>0.5</td>
<td>0.0</td>
<td>RMSE</td>
<td>.11</td>
<td>.11</td>
<td>.12</td>
<td>.15</td>
<td>.11</td>
<td>.22</td>
<td>.20</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bias</td>
<td>.03</td>
<td>.02</td>
<td>.01</td>
<td>-.04</td>
<td>.02</td>
<td>-.01</td>
<td>-.04</td>
<td>-.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IQD</td>
<td>.29</td>
<td>.32</td>
<td>.35</td>
<td>.47</td>
<td>.32</td>
<td>.69</td>
<td>.60</td>
<td>.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UB</td>
<td>.33</td>
<td>.33</td>
<td>.27</td>
<td>.53</td>
<td>.45</td>
<td>.85</td>
<td>.66</td>
<td>.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Coverage</td>
<td>.78</td>
<td>.78</td>
<td>.65</td>
<td>.88</td>
<td>.85</td>
<td>.85</td>
<td>.82</td>
<td>.78</td>
</tr>
<tr>
<td>(4) Lagged dependency</td>
<td>10</td>
<td>0.5</td>
<td>0.5</td>
<td>RMSE</td>
<td>.13</td>
<td>.15</td>
<td>.16</td>
<td>.19</td>
<td>.15</td>
<td>.32</td>
<td>.29</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bias</td>
<td>-.05</td>
<td>-.06</td>
<td>-.06</td>
<td>-.13</td>
<td>-.05</td>
<td>-.01</td>
<td>-.12</td>
<td>-.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IQD</td>
<td>.42</td>
<td>.47</td>
<td>.50</td>
<td>.59</td>
<td>.47</td>
<td>1.07</td>
<td>.88</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UB</td>
<td>.52</td>
<td>.56</td>
<td>.40</td>
<td>.76</td>
<td>.77</td>
<td>1.14</td>
<td>.91</td>
<td>.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Coverage</td>
<td>.85</td>
<td>.84</td>
<td>.72</td>
<td>.89</td>
<td>.89</td>
<td>.81</td>
<td>.76</td>
<td>.68</td>
</tr>
<tr>
<td>(5) Redundant events</td>
<td>10</td>
<td>0.5</td>
<td>0.0</td>
<td>RMSE</td>
<td>.24</td>
<td>.25</td>
<td>.26</td>
<td>.21</td>
<td>.45</td>
<td>.35</td>
<td>.20</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bias</td>
<td>-.17</td>
<td>-.16</td>
<td>-.21</td>
<td>-.10</td>
<td>-.44</td>
<td>-.25</td>
<td>-.18</td>
<td>-.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IQD</td>
<td>.76</td>
<td>.78</td>
<td>.82</td>
<td>.65</td>
<td>.69</td>
<td>.90</td>
<td>.50</td>
<td>.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UB</td>
<td>.82</td>
<td>.55</td>
<td>.84</td>
<td>.98</td>
<td>.82</td>
<td>.93</td>
<td>.63</td>
<td>.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Coverage</td>
<td>.85</td>
<td>.68</td>
<td>.84</td>
<td>.88</td>
<td>.51</td>
<td>.69</td>
<td>.70</td>
<td>.64</td>
</tr>
<tr>
<td>(6) Unobserved events</td>
<td>20</td>
<td>0.5</td>
<td>0.0</td>
<td>RMSE</td>
<td>.15</td>
<td>.16</td>
<td>.18</td>
<td>.16</td>
<td>.31</td>
<td>.29</td>
<td>.13</td>
<td>.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bias</td>
<td>-.05</td>
<td>-.05</td>
<td>-.11</td>
<td>-.03</td>
<td>-.04</td>
<td>-.08</td>
<td>-.08</td>
<td>-.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IQD</td>
<td>.51</td>
<td>.48</td>
<td>.58</td>
<td>.50</td>
<td>.92</td>
<td>.92</td>
<td>.35</td>
<td>.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UB</td>
<td>.59</td>
<td>.45</td>
<td>.86</td>
<td>.78</td>
<td>1.11</td>
<td>.92</td>
<td>.45</td>
<td>.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Coverage</td>
<td>.85</td>
<td>.75</td>
<td>.94</td>
<td>.89</td>
<td>.82</td>
<td>.77</td>
<td>.76</td>
<td>.75</td>
</tr>
</tbody>
</table>

The table shows statistics of standardized IRFs at horizons $h = 0$ and $h = 4$. $m$ is the number of interventions, while $\sigma_\zeta$ is the dispersion of policy shocks and $\omega$ is the weight of the lagged policy target in the policy rule (see equation (9)). RMSE and Bias are the root mean squared error of the estimate and its difference to the true IRF, respectively. IQD is the $[0.1, 0.9]$ interquantile difference of the distribution of the central estimate measuring of its true uncertainty, while UB is the corresponding estimated uncertainty bands. Coverage stands for the share of draws where the true IRF lies within the estimated bands with a desired value of .80. The models and further simulation details, in particular simulations (5) and (6), are explained in the main text. We take 1000 draws.
Table 2: Lagged Dependencies of the Instruments

<table>
<thead>
<tr>
<th></th>
<th>$y_t$</th>
<th>$p_t$</th>
<th>$r_t$</th>
<th>$s_t$</th>
<th>$c^P_t$</th>
<th>$c^H_t$</th>
<th>$p^H_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital requirements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p = 2$</td>
<td>.84</td>
<td>1.69</td>
<td>0.84</td>
<td>1.62</td>
<td>1.12</td>
<td>4.96</td>
<td>4.08</td>
</tr>
<tr>
<td>$p = 4$</td>
<td>4.88</td>
<td>6.02</td>
<td>5.38</td>
<td>6.93</td>
<td>7.16**</td>
<td>13.78</td>
<td>4.20</td>
</tr>
<tr>
<td>Underwriting standards</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p = 2$</td>
<td>1.00</td>
<td>1.65</td>
<td>*7.15</td>
<td>3.56</td>
<td>.02</td>
<td>1.70</td>
<td>1.66</td>
</tr>
<tr>
<td>$p = 4$</td>
<td>3.05</td>
<td>2.09</td>
<td>7.89</td>
<td>*9.51</td>
<td>.84</td>
<td>3.18</td>
<td>*10.03</td>
</tr>
</tbody>
</table>

The table shows the LR statistics of $\beta_{j,1} = \ldots = \beta_{j,p} = 0$ from ordered probits regressing instrument $z_t$ on the variables included in the VAR at lags 1 to $p$. Lags are set to either $p = 2$ or $p = 4$. The statistics are $\chi^2$-distributed with 2 and 4 df, respectively. '*' and '**' indicate significance at 5% and 1% levels, respectively.
Table 3: Average Impact of Policy Innovations

<table>
<thead>
<tr>
<th>Capital requirements</th>
<th>$y_t$</th>
<th>$p_t$</th>
<th>$r_t$</th>
<th>$s_t$</th>
<th>$c^D_t$</th>
<th>$c^H_t$</th>
<th>$p^H_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>32</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>0.1</td>
<td>-.50</td>
<td>-.40</td>
<td>-.40</td>
<td>.01</td>
<td>-1.47</td>
<td>-1.06</td>
<td>-1.39</td>
</tr>
<tr>
<td>0.5</td>
<td>-.26</td>
<td>-.13</td>
<td>-.19</td>
<td>.07</td>
<td>-.66</td>
<td>-.16</td>
<td>-.19</td>
</tr>
<tr>
<td>0.9</td>
<td>-.04</td>
<td>.13</td>
<td>-.00</td>
<td>.13</td>
<td>.07</td>
<td>.55</td>
<td>.99</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Underwriting standards</th>
<th>$y_t$</th>
<th>$p_t$</th>
<th>$r_t$</th>
<th>$s_t$</th>
<th>$c^D_t$</th>
<th>$c^H_t$</th>
<th>$p^H_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>32</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>0.1</td>
<td>-.67</td>
<td>-.26</td>
<td>-.38</td>
<td>.02</td>
<td>-2.06</td>
<td>-2.74</td>
<td>-4.00</td>
</tr>
<tr>
<td>0.5</td>
<td>-.37</td>
<td>-.08</td>
<td>-.17</td>
<td>.10</td>
<td>-.98</td>
<td>-1.31</td>
<td>-2.28</td>
</tr>
<tr>
<td>0.9</td>
<td>-.07</td>
<td>.10</td>
<td>.10</td>
<td>.17</td>
<td>-.05</td>
<td>-.23</td>
<td>-.70</td>
</tr>
</tbody>
</table>

The table shows the median and 0.1 and 0.9 quantiles of responses to the average policy shock from model DSC at different horizons (quarters), as indicated in row $h$.

Table 4: Robustness Check against Lagged Impacts

<table>
<thead>
<tr>
<th>Capital requirements</th>
<th>$y_t$</th>
<th>$p_t$</th>
<th>$r_t$</th>
<th>$s_t$</th>
<th>$c^D_t$</th>
<th>$c^H_t$</th>
<th>$p^H_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>32</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>0.1</td>
<td>-.11</td>
<td>-.15</td>
<td>-.10</td>
<td>-.05</td>
<td>-.27</td>
<td>-.93</td>
<td>-1.11</td>
</tr>
<tr>
<td>0.5</td>
<td>.00</td>
<td>-.01</td>
<td>-.00</td>
<td>-.01</td>
<td>.10</td>
<td>-.30</td>
<td>-.21</td>
</tr>
<tr>
<td>0.9</td>
<td>.13</td>
<td>.11</td>
<td>.10</td>
<td>.02</td>
<td>.57</td>
<td>.12</td>
<td>.29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Underwriting standards</th>
<th>$y_t$</th>
<th>$p_t$</th>
<th>$r_t$</th>
<th>$s_t$</th>
<th>$c^D_t$</th>
<th>$c^H_t$</th>
<th>$p^H_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>32</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>0.1</td>
<td>-.28</td>
<td>-.16</td>
<td>-.13</td>
<td>-.02</td>
<td>-.53</td>
<td>-.85</td>
<td>-.90</td>
</tr>
<tr>
<td>0.5</td>
<td>-.08</td>
<td>-.01</td>
<td>-.00</td>
<td>.01</td>
<td>.09</td>
<td>-.19</td>
<td>-.09</td>
</tr>
<tr>
<td>0.9</td>
<td>.03</td>
<td>.14</td>
<td>.11</td>
<td>.04</td>
<td>.26</td>
<td>.23</td>
<td>.46</td>
</tr>
</tbody>
</table>

The table shows the median and 0.1 and 0.9 quantiles of lagged responses to policy interventions from model DSC at different horizons (quarters), as indicated in row $h$. 

26
Figure 1: Beta-Binomial Priors for Sign Concordance

The graphs show densities \( f(\varphi; m, \Lambda, p, q) \) of beta-binomial distributions with the number of events set to \( m = 10 \) and \( m = 50 \). The beta distribution is defined over support \([\Lambda, 1]\) with parameters \( p \) and \( q \).

Figure 2: Sign Concordance Posterior Densities

The shaded area shows the posterior density of the sign concordance statistics \( m\varphi \) for model \( DSC \). The lines show the same posterior density for models \( DC \) and \( SC \) and the sign concordance prior.
The graphs show the median estimates of IRFs to a shock of 1% from models $DSC$, $DC$, and $SC$, together with $[0.10, 0.90]$ quantiles for model $DSC$. 
Supporting Information For Online Publication
Supplement A: The Narrative Indicators

Capital Requirements

1981/15/12  Tightening

The Federal Reserve Board and the Office of the Comptroller of the Currency introduce capital standards common to all banks. The standards employ a leverage ratio of primary capital (which consisted mainly of equity and loan loss reserves) to average total assets. Standards differ slightly by type of institution with a value of 6% for community banks and 5% for large regional institutions. Source: Federal Deposit Insurance Corporation (FDIC).

1983/03/01  Tightening

Congress passes the International Lending Supervision Act (ILSA). This statute directs the banking regulators to “achieve and maintain adequate capital by establishing minimum levels of capital” for banks subject to regulation. The ILSA is enacted in response to the Latin American debt crisis, which revealed a high risk of the foreign sovereign debt exposure of some U.S. banks. The law also put on firmer footing the regulators’ authority to issue capital adequacy rules. Source: Federal Register.

1985/15/06  Tightening

Regulators abolish the differences in bank leverage by type of bank as established in the 1981/15/12 Act in favor of a uniform standard of 5.5%. Banks with less than 3% of primary-capital-to-total assets are declared to be “operating in unsafe condition” and are made subject to enforcement actions. Source: FDIC.

1990/31/12  Tightening

The first stage of the Basel I rules is enacted by US regulators imposing two requirements on capital ratios, related to Tier 1 and Tier 2 capital. First, Basel I calls for a minimum ratio of total (Tier 1 plus Tier 2) capital to risk-weighted assets (RWA) of 8%, and of Tier 1 capital to risk-weighted assets of 4%. The first stage requires respective ratios of 7.25% and 3%, while the full are phased in until the end of 1992. Source: Posner (2014).

1991/19/12  Tightening

The Federal Deposit Insurance Corporation Improvement Act categorizes institutions according to their capital ratios. Other than “well capitalized” banks (at least 10% total risk-based, 6% Tier 1 risk-based, and 5% leverage capital ratios) face restrictions on certain activities and are subject to mandatory or discretionary supervisory actions. Source: Government Publishing Office (GPO).

1992/31/12  Tightening

The final implementation stage of the Basel I rules is enacted by US regulators with the own funds ratio set to 8%, and the leverage ratio set to 4%. Source: Posner (2014).
2002/01/01  Easing

The Recourse Rule reduces risk weights for AAA- and AA- rated “private-label” mortgage-backed securities (MBS) and collateralized debt obligation (CDO) tranches originated by large banks to 0.2 in line with government-sponsored enterprise (GSE)–originated MBS. For A-rated tranches, the risk weights are set to 0.5, while lower-rated tranches are assigned higher risk weights. The rule is designed to encourage securitization without encouraging risk taking, while risk weights are kept close to the 2004 Basel II risk weights. Source: Posner (2014).

2006/31/12  Tightening

The Tier 1 leverage ratio is increased to 4 %. Source: Posner (2014).

2013/01/01  Tightening

The Federal Reserve Board approves a final rule to implement changes to the market risk capital rule, which requires banking organizations with significant trading activities to adjust their capital requirements to better account for the market risks of those activities (Basel II.5). The adoption of Basel II.5, also known as the market capital risk rule, has been issued by the U.S. federal banking regulators on June 7, 2012. Source: Federal Reserve Board (FRB).

2013/30/07  Tightening

The Federal Reserve Board (FRB) introduces a supplementary leverage ratio requirement of 3% for banks using the advanced approach for RWA calculation. An additional 2% buffer requirement has been proposed for G-SIBs. Further, IRB banks are required to apply the lower of the capital ratios calculated under the standardized and IRB approaches. Source: FRB.

Mortgage Underwriting Standards

1958/01/04  Easing

Changes to requirements on loans insured by the Veteran Administration. Removal of 2% downpayment requirement on insured loans. Act of Congress changes requirements on loans insured by the Federal Housing Administration. (i) LTV for new construction, 97% of first $13,500 of value plus 85% of next USD 2,500 plus 70% of value in excess of $16,000 to maximum mortgage of USD 20,000. (ii) LTV for existing construction, 90% of first US$D 13,500 of value plus 85% of next $ 2,500 plus 70% of value in excess of $ 16,000 to maximum mortgage of $ 20,000. Source: Elliot et al (2013).

1959/23/09  Easing

Act of Congress changes requirements on loans insured by the Federal Housing Administration. (i) LTV for new construction, 97% of first $13,500 of value plus 90% of next $4,500 plus 70% of value in excess of $18,000 to maximum mortgage of $22,500. (ii) LTV for existing construction, 90% of first $18,000 of value plus 70% of value in excess of $18,000 to maximum mortgage of $22,500. Source: Elliot et al (2013).
1961/30/06  Easing

Act of Congress changes requirements on loans insured by the Federal Housing Administration. (i) LTV for new construction set to 97% of first $15,000 of value plus 90% of next $5,000 plus 75% of value in excess of $20,000 to maximum mortgage of $25,000. (ii) LTV for existing construction, 90% of first $20,000 of value plus 75% of value in excess of $20,000 to maximum mortgage of $25,000. (iii) Easing of maturity standards for new construction, maximum mortgage term raised from 30 to 35 years or 3/4 of the remaining life of improvements, whichever is less; existing construction still 30 years. Source: Elliot et al (2013).

1964/01/01  Easing

National banks are allowed to extend real estate loans with 25-year terms and 80% LTV if fully amortized. Source: Elliot et al (2013).

1964/02/09  Easing

Act of Congress changes requirements on loans insured by the Federal Housing Administration. (i) LTV for new construction, 97% of first $15,000 of value plus 90% of next $5,000 plus 75% of value in excess of $20,000 to maximum mortgage of $30,000. (ii) LTV for existing construction, 90% of first $20,000 of value plus 75% of value in excess of $20,000 to maximum mortgage of $30,000. Source: Elliot et al (2014).

1965/10/08  Easing

Act of Congress changes requirements on loans insured by the Federal Housing Administration. (i) LTV for new construction, 97% of first $15,000 of value plus 90% of next $5,000 plus 80% of value in excess of $20,000 to maximum mortgage of $30,000. (ii) LTV for existing construction, 90% of first $20,000 of value plus 80% of value in excess of $20,000 to maximum mortgage of $30,000. Source: Elliot et al (2013).

1970/01/01  Easing

National banks are allowed to extend real estate loans with 30-year terms and 90% LTV if fully amortized. Source: Elliot et al (2013).

1974/01/01  Easing

National banks are allowed to extend real estate loans with 30-year terms and 90% LTV if 75% amortized. Source: Elliot et al (2013).

1983/01/09  Easing

LTV limits are removed for all bank mortgage loans (Garn-St Germain). Source: Elliot et al (2013).

2014/30/01  Tightening

A New Ability to Repay (ATR) and Qualified Mortgage (QM) Rule by Consumer Financial Protection Bureau (CFPB) establishes a minimum set of underwriting standards in the mortgage market. For qualified mortgages the borrower must prove a debt service-to-income ratio no greater than 43%. Source: CFPB.
Supplement B.1: Linear Discriminant Analysis

This supplement outlines the relation of the DC regression to discriminant analysis, following Maddala (2013). Consider a dichotomous variable $z_t$ that takes the value $z_t = 1$ for $m$ observations and $z_t = 0$ for the remaining $T - m$ observations. The objective of discriminant analysis is to estimate function $\psi(x_t)$ to predict $z_t$ from a set of random variables $x_t = (x_{1,t}, \ldots, x_{n,t})$ based on the rule $\hat{z}_t = 1$ if $\psi(x_t) > 0$ and $\hat{z}_t = 0$ otherwise (e.g. Maddala, 2013: 79ff). $\psi(x_t)$ is chosen to minimize the objective function

$$ C = C_1 \int_{R_1} f_1(x_t)dx + C_0 \int_{R_0} f_0(x_t)dx, $$

where $f_k(x_t)$ denote the conditional distributions of $x_t|z_t = k$. $R_1$ defines the region such that $\psi(x_t) > 0$ if $x_t \in R_1$ and $R_0$ is the complement of $R_1$. $C_k$ is the cost of misclassifying a member of group $G_k$.

Under the assumption that $x_t|z_t = 1 \sim N(\mu_1, \Sigma)$ and $x_t|z_t = 0 \sim N(\mu_0, \Sigma)$, the optimal discriminant function is linear, $\psi(x_t) = \psi_1^T x_t$. Under the specific loss function $mC_1 = (T - m)C_0$, the maximum likelihood estimate of parameter vector $\psi_1$ maximizes the ratio of the squared difference in means between groups and the variance within groups, $(\psi_1^T \Sigma \psi_1)^{-1} [\psi_1^T (\mu_1 - \mu_0)]^2$. This is equivalent up to scale to estimating $\alpha$ via OLS from the regression $z_t^* = a_0 + a^T x_t$, where $z_t^* = z_t - m/T$ (Maddala, 2013:18ff).
Supplement B.2: Further Details on the Structural VAR

Consider the moving average representation of equation (2)

\[ x_t = \left( \sum_{s=0}^{\infty} \Psi_s \right) A_0^{-1} c + \sum_{s=0}^{\infty} \Psi_s A_0^{-1} \varepsilon_{t-s} \]

where matrices \( \Psi_s \) are the elements of lag polynomial \( \Psi(L) = B^{-1}(L) \) with \( B(L) = I_n - \sum_{s=1}^{p} B_s L^s \). \( \Psi(L) \) defines the IRF of SVAR given by equations (1) and (2). Since \( \Psi_0 = I_n \), matrix \( A_0^{-1} \) gives the contemporaneous impact of the structural innovations on the VAR series.

**Constructing matrix \( A_0 \)**

We first review the construction of matrix \( A_0 \) as e.g. set out in Arias et al. (2018). The condition \( u_t = A_0^{-1} \varepsilon_t \), together with \( E u_t u_t^T = \Sigma \) and \( E \varepsilon_t \varepsilon_t^T = \Sigma \), implies \( \Sigma^{-1} = A_0^T A_0 \). Further, matrix \( A_0 \) can be expressed as \( A_0^T = A_r Q \), where \( A_r \) is a unique lower triangular matrix derived from the Choleski decomposition \( \Sigma^{-1} = A_r A_r^T \) and \( Q = (q_1, \ldots, q_n) \) is an arbitrary orthogonal matrix, \( Q^T = Q^{-1} \), that is constructed such that \( A_0 \) satisfies certain restrictions. Arias et al. (2018) show how random draws of \( Q \) that satisfy deterministic restrictions may be constructed in a recursive way from a Gram-Schmidt orthogonalisation (GSO): column \( q_j \) is obtained by drawing an \( n \times 1 \) vector \( x_j \sim N(0, I_n) \) and deriving \( q_j \) along the lines of the GSO such that \( q_j \) is orthogonal to \((q_k, \ldots, q_{j-1})\) and satisfies further deterministic restrictions specific to innovations \( \varepsilon_{t,j} \).

In case of the DC restriction, vector \( \alpha \) defines the first column of \( A_0 \), which implies \( q_1 = A_r^{-1} \alpha \). The reverse expression \( \alpha = A_r q_1 \) is used case of the SC restriction. An uninformative random draw of \( \alpha \) is obtained by drawing \( q_1 \) from the Haar measure of orthogonal matrices as \( q_1 = v/||v|| \), with random draw \( v \sim N(0, I_n) \).

In both cases, the remaining columns of matrix \( Q \) are irrelevant and are constructed from application of the GSO without further restrictions as explained in Arias et al. (2018). Note that \( q_1 \) suffices for defining the contemporaneous impact of \( \theta_1 \), as \( A_0^{-1} = (A_r^T)^{-1} Q \) and the first column of \( A_0^{-1} \) is therefore well-defined and independent of all \( q_j \) with \( j > 1 \).

**Sampling**

This details the sampling of \( \gamma \) and the calculation of mean shift \( \Delta \) described in section 2.3. The assumption

\[ \theta_1 | z_t \sim N(\gamma z_t, \sigma_z) \]
\[ \varepsilon_t^\top | z_t \sim N(0, I_{n-1}) \]
implies that $\hat{\gamma} = m^{-1} \sum_{t=1}^{T} \theta_t z_t$ is subject to standard inference conditional on $B^+$ and $\alpha$. Assuming an uninformative prior for $\gamma$ and a Jeffrey prior for $\sigma^2_\gamma$ gives the expressions as in the main text.

The expected value of the VAR forecast errors conditional on $z_t \neq 0$ is then given as

$$\mu_1 = \mathbb{E}(u_t | z_t = 1) = A_0^{-1} \begin{pmatrix} \gamma \\ 0_{n-1} \end{pmatrix}$$

Given $\mathbb{E}u_t = 0$ this implies

$$0 = (m^+ - m^-) \mu_1 + (T - m)^{-1} \mathbb{E}(u_t | z_t = 0)$$

$$\mathbb{E}(u_t | z_t = 0) = \frac{m^+ - m^-}{T - m} \mu_1.$$

where $m^+$ and $m^-$ are the numbers of positive and negative entries in $z_t$ and $m = m^+ + m^-$. The expression for $\Delta$ in section 2.3 follows from $\Delta = \mu_1 - \mathbb{E}(u^t | z_t = 0)$.

**Combination with Sign and Zero Restrictions**

$DC$ and $SC$ restrictions may also be embedded in the approach of Arias et al. (2018) and thereby be combined with zero and sign restrictions on IRFs. Define $g(A_0, \Psi(L)) = [\Psi^T_0, \Psi^T_1, \ldots, \Psi^T_s] A_0^{-1}$. Express zero and sign restrictions on column $j$ of $\Psi(L)$, i.e. the IRFs to shock $\epsilon_{j,t}$ as

$$Z_j g(A_0, \Psi(L)) e_j = 0$$

$$S_j g(A_0, \Psi(L)) e_j > 0$$

with appropriate selection matrices $Z_j$ and $S_j$. Vector $e_j$ denotes column $j$ of identity matrix $I_n$.

The algorithm of Arias et al. (2018) to generate posterior draws of $\Psi(L) A_0^{-1}$ under these type of restrictions proceeds by (i) drawing from the posterior $(B(L), \Sigma)$ to obtain $\Psi(L)$ and $A_\ast$; (ii) obtaining uninformative draws of $Q$ that satisfy the zero restrictions $Z_j g(A, \Psi(L))$; (iii) applying an importance sampling step to account for volume changes due to zero restrictions; and (iv) inspecting the validity of sign restrictions.

With the $DC$ regression, the draw of $\alpha$ uniquely defines $q_1 = A_\ast^{-1} \alpha$, while the remaining columns of $Q$ remain unspecified. The $DC$ restriction may therefore be combined with zero and sign restrictions on shocks $\epsilon_{t,j}$ for $j > 1$. Note that we draw $\alpha$ from a non-degenerate distribution. Hence, there is no volume reduction and the importance sampling step by Arias et al. (2018) is not required. The $SC$ posterior on shock $\theta_t$ is implemented from a rejection sampling step. Hence, it may be combined with sign restrictions on shocks $\epsilon_{t,j}$ for all $j$ and zero restrictions for $j > 1$. 
Supplement B.3: Monte Carlo Simulation Data Generating Process

We set

\[ B_1 = \rho \begin{bmatrix} \cos(\omega) & \sin(\psi) \\ -\sin(\psi) & \cos(\psi) \end{bmatrix} \quad A_0^{-1} = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1.0 \end{bmatrix}^{1/2} \begin{bmatrix} \cos(a) & \sin(a) \\ -\sin(a) & \cos(a) \end{bmatrix} \]

with \( \rho = 0.9 \), \( \psi = 0.2 \), and \( a = \pi/4 \). Matrix \( B_1 \) is subject to complex conjugate roots and generates cyclical fluctuations of length of \( 2\pi/\psi = 32 \) quarters. Matrix \( A_0^{-1} \) is constructed from a Choleski decomposition times a rotation matrix, such that residuals \( u_t = A_0^{-1} \xi_t \), where \( \xi_t \sim N(0, 0.01I_2) \), are subject to a correlation of close to 0.3, while the rotation matrix ensures that the IRF of \( x_{t,1} \) to shock 1 has the desired shape.

To calibrate the number of policy shocks \( m \) we let \( \sigma_{\nu} = 0.01 \) and calibrate the expected value of the size of the policy innovation, \( \zeta \), to achieve the desired expected number of policy interventions \( m \). This give values of \( \zeta = 1.64 \) for \( m = 10 \) and \( \zeta = 1.28 \) for \( m = 20 \). As regards simulation (4), since \( \text{var} x_{t,i} = (1 - \rho^2)^{-1}\Sigma \), with \( \omega = 0.5 \), the lagged term \( x_{2,t-1} \) explains about 70% of the total variance of \( \zeta^* \). The results presented are based on 1000 draws of the DGP (9) and, for each draw of the DGP, 200 draws of the posterior or bootstrap confidence bounds, respectively. The number of observations is set to \( T = 200 \).

The Bayesian proxy and narrative VARS are explained in the main text. In all cases, we employ an uninformative Jeffrey prior for the reduced form VAR and simply assume 1 lag in estimation, mirroring the data generating process. For model \( BV \zeta \) we then draw from proxy regression \( \zeta_t = a^T u_t + \xi_t \) using an uninformative Normal-Gamma prior as with the \( DC \) regression. In implementing the frequentist proxy VAR, we rely on the code of Mertens and Montiel-Olea (2018), which offer the bootstraps of Jentsch and Lunsford (2016) and Montiel-Olea et al. (2022) for proxy VARs with sparse instruments. The results from the two bootstraps are very similar and we therefore report only the latter.

the final two models use \( z_t \) as a regressor instead of an instrument. For local projections we estimate the equation \( x_{1,t+h} = a^T x_t + \gamma_h z_t + u_{1,t} \) and obtain the impulse response to the policy innovation at horizon \( h \) directly from coefficient \( \gamma_h \). The recursive VAR proposed by Plagborg-Møller and Wolf (2021) amounts to estimating the system \( y_t = B_1 y_{t-1} + A_0 y_t + u_t \), where \( y_t = (z_t, x^T_t) \) and \( A_0 \) is a lower triangular matrix with a zero diagonal. However, with a binary sparse regressor, we obtain better results with setting all coefficients in the first equation for \( z_t \) to zero. Hence, the recursive VAR is equivalent to a VAR for \( x_t \), where \( z_t \) and \( z_{t-1} \) are added as deterministic dummy regressors. For both methods we obtain uncertainty bands from standard bootstraps.
Supplement B.4: Banking Deregulation Index

Our banking deregulation index is an unweighted average of two sub-indices related to inter-state and intra-state deregulation. Each sub-index takes values of zero (full regulation) to one (no regulation) with intermittent values equal to the GDP shares (as of 1980) of states, which had introduced respective deregulation. Hence, the index equals zero before 1970, the beginning of deregulation, and one after 1996.

As discussed by Kroszner and Strahan (1999, 2014), deregulation was a gradual process that consolidated the fragmented banking system in multiple ways. States differed in the timing of when they allowed banks from other states to operate in their jurisdiction and in how many other states were given access. Another source of variation was the timing of the removal of intra-state branching restrictions that prohibited banks to expand their branch network within a state.¹⁴

Figure B.1: Banking Deregulation Index

We use the indices provided by Mian et al (2017), which reflect the start of the deregulation process. For example, the year of inter-state banking deregulation is defined as the first year in which a state allowed out-of-state banks to open a branch. These decisions were based on bilateral arrangements between states, until the Riegle-Neal Act of 1994 resulted in a general deregulation of U.S. inter-state banking. Kroszner and Strahan (1999, 2014) conclude that the process of deregulation was largely exogenous to macro-economic conditions as it was driven by a combination of technological change and shifts in private and public interest. For instance, the speed of deregulation is highly correlated with republican versus democratic state government.

The black solid line shows the true IRF. The blue solid and dotted lines show the central estimate and its [.10, .90] quantiles as provided by the various methods. The shaded area shows the [.10, .90] quantiles of confidence bounds. See Table 1 for the definition of the simulations. The models and the calculation of central estimates and confidence bounds are explained in the main text.
Figure C.2: Standardized IRFs for DC and SC Restrictions

The graphs show the impulse responses to a 1% shock based on either DC or SC restrictions. The solid line shows the median and bounds show [0.05; 0.95] and [0.16; 0.84] quantiles of IRFs. The dotted line shows the main estimate from the DSC restriction.
Figure C.3: IRFs Scaled by the Average Policy Impact

The graphs show the IRFs scaled the impact of average policy shock of size based on either \textit{DC} or \textit{SC} restrictions. See Figure A.3 for further explanations.
Figure C.4: Robustness Check against Lagged Impacts

The graphs show standardised IRFs from the robustness checks against lagged innovations discussed in section 4.3. The left hand graph shows estimates with lags drawn for individual periods from a uniform distribution (model DSC) and using the SC prior, the right hand one for fixed lags from 1 to 4 (model DC). The latter corresponds to the results shown in Table 4. See Figure A.3 for further explanations.
Figure C.5: Standardized IRFs from Alternative Estimates

The left hand graph shows estimates including the deregulation index. The right hand graph shows estimates from a VAR including 4 lags. See Figure A.3 for further explanations.
Figure C.6: Impulse Responses from Proxy and Recursive VARs

The graphs show IRFs from the proxy VAR with confidence bounds from the bootstrap by Montiel Olea et al. (2020) and from the recursive VAR by Plagborg-Møller and Wolf (2021) with standard bootstrap bounds. See Figure A.3 for further explanations.