

## IDENTIFICATION AT THE ZERO LOWER BOUND

SOPHOCLES MAVROEIDIS

Department of Economics, University of Oxford

I show that the zero lower bound (ZLB) on interest rates can be used to identify the causal effects of monetary policy. Identification depends on the extent to which the ZLB limits the efficacy of monetary policy. I propose a simple way to test the efficacy of unconventional policies, modeled via a “shadow rate.” I apply this method to U.S. monetary policy using a three-equation structural vector autoregressive model of inflation, unemployment, and the Federal Funds rate. I reject the null hypothesis that unconventional monetary policy has no effect at the ZLB, but find some evidence that it is not as effective as conventional monetary policy.

KEYWORDS: SVAR, censoring, coherency, partial identification, monetary policy, shadow rate.

### 1. INTRODUCTION

THE ZERO LOWER BOUND (ZLB) on nominal interest rates has arguably been a challenge for policy makers and researchers of monetary policy. Policy makers have had to resort to so-called unconventional policies, such as quantitative easing or forward guidance, which had previously been largely untested. Researchers have to use new theoretical and empirical methodologies to analyze macroeconomic models when the ZLB binds. So, the ZLB is generally viewed as a problem or at least a nuisance. This paper proposes to turn this problem on its head to solve another long-standing question in macroeconomics: the identification of the causal effects of monetary policy on the economy.

The intuition is as follows. If the ZLB limits the ability of policy makers to react to macroeconomic shocks, as argued, for example, by Eggertsson and Woodford (2003), the response of the economy to shocks will change when the policy instrument hits the ZLB. Because the difference in the behavior of macroeconomic variables across the ZLB and non-ZLB regimes is only due to the impact of monetary policy, the switch across regimes provides information about the causal effects of policy. In the extreme case that monetary policy completely shuts down during the ZLB regime, either because policy makers do not use alternative (unconventional) policy instruments, or because such instruments turn out to be completely ineffective, the only difference in the behavior of the economy across regimes is due to the impact of (conventional) policy during the unconstrained regime. Therefore, the ZLB identifies the causal effect of policy during the unconstrained regime. If monetary policy remains partially effective during the ZLB regime, for example, through the use of unconventional policy instruments, then the difference in the behavior of the economy across regimes will depend on the difference in the effectiveness

---

Sophocles Mavroeidis: [sophocles.mavroeidis@economics.ox.ac.uk](mailto:sophocles.mavroeidis@economics.ox.ac.uk)

This research is funded by the European Research Council via Consolidator Grant 647152. I would like to thank Guido Ascari, Sergio de Ferra, James Duffy, Andrea Ferrero, Jim Hamilton, Daisuke Ikeda, Federica Romei, Frank Schorheide, Francesco Zanetti, and seminar participants at the Chicago Fed, New York Fed, Board of Governors, Bank of Japan, Hitotsubashi University, Keio University, UCL, University of Cambridge, University of Oxford, University of Warwick, DNB, University of Pennsylvania, Pennsylvania State University, NBER Summer Institute, IAAE conference, Netherlands Econometrics Study Group Meeting, World Congress and North American Winter meetings of the Econometric Society, Athens University of Economics and Business, and University of Cyprus for useful comments and discussion, as well as Mishel Ghassibe, Lukas Freund, Shangshang Li, and Patrick Vu for research assistance.

© 2021 The Author. *Econometrica* published by John Wiley & Sons Ltd on behalf of The Econometric Society. Sophocles Mavroeidis is the corresponding author on this paper. This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

of conventional and unconventional policies. In this case, we obtain only partial identification of the causal effects of monetary policy, but we can still get informative bounds on the relative efficacy of unconventional policy. In the other extreme case that unconventional policy is as effective as conventional policy, there is no difference in the behavior of the economy across regimes, and we have no additional information to identify the causal effects of policy. However, we can still test this so-called ZLB irrelevance hypothesis (De-bortoli, Gali, and Gambetti (2019)) by testing whether the reaction of the economy to shocks is the same across the two regimes.

There are similarities between identification via occasionally binding constraints and identification through heteroscedasticity (Rigobon (2003)), or more generally, identification via structural change (Magnusson and Mavroeidis (2014)). That literature showed that the switch between different regimes generates variation in the data that identifies parameters that are constant across regimes. For example, an exogenous shift in a policy reaction function or in the volatility of shocks identifies the transmission mechanism, provided the latter is unaffected by the policy shift. When the switch from one regime to another is exogenous, regime indicators are valid instruments, and the methodology in Magnusson and Mavroeidis (2014) is applicable. However, regimes induced by occasionally binding constraints are not exogenous—whether the ZLB binds or not clearly depends on the structural shocks, so regime indicators cannot be used as instruments in the usual way, and a new methodology is needed to analyze these models.

In this paper, I show how to control for the endogeneity in regime selection and obtain identification in structural vector autoregressions (SVARs).<sup>1</sup> The methodology is parametric and likelihood-based, and the analysis is similar to the well-known Tobit model (Tobin (1958)). More specifically, the methodological framework builds on the early microeconometrics literature on simultaneous equations models with censored dependent variables (see Amemiya (1974), Lee (1976), Blundell and Smith (1994)), and the more recent literature on dynamic Tobit models (see Lee (1999)) and particle filtering (see Pitt and Shephard (1999)).

A further contribution of this paper is a general methodology to estimate reduced-form VARs with a variable subject to an occasionally binding constraint. This is a necessary starting point for SVAR analysis that uses any of the existing popular identification schemes, such as short- or long-run restrictions, sign restrictions, or external instruments. In the absence of any constraints, reduced-form VARs can be estimated consistently by ordinary least squares (OLS), which is Gaussian maximum likelihood, or its corresponding Bayesian counterpart, and inference is fairly well-established. However, it is well known that OLS estimation is inconsistent when the data are subject to censoring or truncation; see, for example, Greene (1993) for a textbook treatment. So, it is not possible to estimate a VAR consistently by OLS using any sample that includes the ZLB, or even using (truncated) subsamples when the ZLB is not binding (because of selection bias), as was pointed out by Hayashi and Koeda (2019). It is not possible to impose the ZLB constraint using Markov switching models with exogenous regimes, as in Liu, Theodoridis, Mumtaz, and Zanetti (2019), because exogenous Markov switching cannot guarantee that the constraint will be respected with probability 1, and also does not account for the

---

<sup>1</sup>There is a related literature on dynamic stochastic general equilibrium (DSGE) models with a ZLB; see, for example, Fernández-Villaverde, Gordon, Guerrón-Quintana, and Rubio-Ramírez (2015), Guerrieri and Iacoviello (2015), Aruoba, Cuba-Borda, and Schorfheide (2017), Kulish, Morley, and Robinson (2017), and Aruoba, Cuba-Borda, Higa-Flores, Schorfheide, and Villalvazo (2020). The papers in this literature do not point out the implications of the ZLB for identification of monetary policy shocks.

fact that the switch from one regime to the other depends on the structural shocks. Finally, it is not possible to perform consistent estimation and valid inference on the VAR (i.e., error bands with correct coverage on impulse responses) using externally obtained measures of the shadow rate, such as the one proposed by [Wu and Xia \(2016\)](#), as any such measures are subject to large and persistent estimation error that is not accounted for if they are treated as known in subsequent analysis. See also [Rossi \(2019\)](#) for a comprehensive discussion of the challenges posed by the ZLB for the estimation of structural VARs.

The methodology developed in this paper allows for the presence of a shadow rate, estimates of which can be obtained, but more importantly, it fully accounts for the impact of sampling uncertainty in the estimation of the shadow rate on inference about the structural parameters such as impulse responses. Therefore, the paper fills an important gap in the literature, as it provides the requisite methodology to implement any of the existing identification schemes. [Hayashi and Koeda \(2019\)](#) developed a VAR model with endogenous regime switching in which the policy variables that are subject to a lower bound are modeled using Tobit regressions. A key difference of their methodology from the one developed here is that they imposed recursive identification of monetary policy shocks, which the present paper shows to be an overidentifying, and hence testable, restriction. Moreover, their model does not include shadow rates. A more recent paper by [Aruoba, Schorfheide, and Villalvazo \(2020\)](#) also studies SVARs with occasionally binding constraints, but does not focus on the implications of these constraints for identification.

Identification of the causal effects of policy by the ZLB does not require that the policy reaction function be stable across regimes. However, inference on the efficacy of unconventional policy, or equivalently, the causal effects of shocks to the shadow rate over the ZLB period, obviously depends on whether or not the reaction function remains the same across regimes. For example, an attenuation of the causal effects of policy over the ZLB period may indicate that unconventional policy is only partially effective, but it is also consistent with unconventional policy being less active (during ZLB regimes) than conventional policy (during non-ZLB regimes). This is a fundamental identification problem that is difficult to overcome without additional information, such as measures of unconventional policy stance, or additional identifying assumptions, such as parametric restrictions or external instruments. This can be done using the methodology developed in this paper.

The structure of the paper is as follows. Section 2 presents the main identification results of the paper in the context of a static bivariate simultaneous equations model with a limited dependent variable subject to a lower bound. Section 3 generalizes the analysis to a SVAR with an occasionally binding constraint and discusses identification, estimation, and inference. Section 4 provides an application to a three-equation SVAR in inflation, unemployment, and the Federal Funds rate from [Stock and Watson \(2001\)](#). Using a sample of post-1960 quarterly U.S. data, I find some evidence that the ZLB is empirically relevant, and that unconventional policy is only partially effective. Proofs and simulation results are given in the [Appendix](#) at the end. Additional supporting material is provided in the Supplemental Material ([Mavroeidis \(2021\)](#)) available online.

## 2. SIMULTANEOUS EQUATIONS MODEL

I first illustrate the idea using a simple bivariate simultaneous equations model (SEM), which is both analytically tractable and provides a direct link to the related microeconomics literature. To make the connection to the leading application, I will motivate this

using a very stylized economy without dynamics in which the only outcome variable is inflation  $\pi_t$  and the (conventional) policy instrument is the short-term nominal interest rate,  $r_t$ . In addition to the traditional interest rate channel, the model allows for an “unconventional monetary policy” channel that can be used when the conventional policy instrument hits the ZLB. An example of such a policy is quantitative easing (QE), in the form of long-term asset purchases by the central bank. Here I discuss a simple model of QE.<sup>2</sup>

Abstracting from dynamics and other variables, the equation that links inflation to monetary policy is given by

$$\pi_t = c + \beta(r_t - r^n) + \varphi b_{L,t} + \varepsilon_{1t}, \tag{1}$$

where  $c$  is a constant,  $r^n$  is the neutral rate,  $b_{L,t}$  is the amount of long-term bonds held by the private sector in log-deviation from its steady state, and  $\varepsilon_{1t}$  is an exogenous structural shock unrelated to monetary policy. Equation (1) can be obtained from a model of bond-market segmentation, as in [Chen, Cúrdia, and Ferrero \(2012\)](#), where a fraction of households is constrained to invest only in long-term bonds see the Supplemental Material for more details. In such a model, the parameter  $\varphi$  that determines the effectiveness of QE is proportional to the fraction of constrained households and the elasticity of the term premium with respect to asset holdings, both of which are assumed to be outside the control of the central bank.

The nominal interest rate is set by a Taylor rule subject to the ZLB constraint, namely,

$$r_t = \max(r_t^*, 0), \tag{2a}$$

$$r_t^* = r^n + \gamma \pi_t + \varepsilon_{2t}, \tag{2b}$$

where  $r_t^*$  represents the desired target policy rate, and  $\varepsilon_{2t}$  is a monetary policy shock. When  $r_t^*$  is negative, it is unobserved. The unobserved  $r_t^*$  will be referred to as the “shadow rate,” and it represents the desired policy stance prescribed by the Taylor rule in the absence of a binding ZLB constraint.

Suppose that QE is activated only when the conventional policy instrument  $r_t$  hits the ZLB,<sup>3</sup> and follows the same policy rule (2b), up to a factor of proportionality  $\alpha$ , that is,

$$b_{L,t} = \min(\alpha r_t^*, 0).$$

Substituting for  $b_{L,t}$  in equation (1) and letting  $\beta^* := \alpha\varphi$ , we obtain

$$\pi_t = c + \beta(r_t - r^n) + \beta^* \min(r_t^*, 0) + \varepsilon_{1t}. \tag{3}$$

A special case arises when QE is ineffective ( $\varphi = 0$ ), or the monetary authority does not pursue a QE policy ( $\alpha = 0$ ), so that equation (3) becomes

$$\pi_t = c + \beta(r_t - r^n) + \varepsilon_{1t}, \tag{4}$$

and monetary policy is completely inactive at the ZLB.

<sup>2</sup>The more general SVAR model of the next section can also incorporate forward guidance in the form of [Reifschneider and Williams \(2000\)](#) and [Debortoli, Gali, and Gambetti \(2019\)](#), as shown in [Ikeda, Li, Mavroeidis, and Zanetti \(2020\)](#).

<sup>3</sup>The assumption that QE is only active during the ZLB regime is only made for simplicity, as it is inconsequential for the resulting functional form of the transmission equation. We can let QE be active all the time, and even allow for a different rule for QE above and below the ZLB, that is,  $b_{L,t} = \alpha \min(r_t^*, 0) + \alpha_1 \max(r_t^*, 0)$ . Then, substituting back into (1) yields an equation that is isomorphic to (3), that is,  $\pi_t = c_1 + \beta r_t + \tilde{\beta}^* \min(r_t^*, 0) + \varepsilon_{1t}$ , with  $\tilde{\beta} := \beta + \varphi\alpha_1$  and  $\tilde{\beta}^* := \varphi\alpha$ .

Another special case of the model given by equations (2) and (3) arises when  $\beta^* = \beta$  in equation (3). This happens when  $\varphi \neq 0$  and  $\alpha$  is chosen by the monetary authority to be equal to  $\beta/\varphi$ . This can be done when policy makers know the transmission mechanism in equation (1) and have no restrictions in setting the policy parameter  $\alpha$  so as to fully remove the impact of the ZLB on conventional policy. In that case, the equation for the outcome variable becomes

$$\pi_t = c + \beta(r_t^* - r^n) + \varepsilon_{1t}. \quad (5)$$

The model given by equations (2) and (5) is one in which monetary policy is completely unconstrained and there is no difference in outcomes across policy regimes. Such models have been put forward by Swanson and Williams (2014), Debortoli, Gali, and Gambetti (2019), and Wu and Zhang (2019).

The nesting model given by equation (3) allows the effects of conventional and unconventional policy to differ. This could reflect informational as well as political or institutional constraints that prevent policy makers from calibrating their unconventional policy response to match exactly the policy prescribed by the Taylor rule. For instance, it may be that policy makers do not know the effectiveness of the QE channel  $\varphi$ , or that the scale of asset purchases needed to achieve the desired policy stance during a ZLB regime is too large to be politically acceptable. Such a consideration may be particularly pertinent, for example, in the Eurozone. Importantly, one does not need to take a theoretical stand on this issue, because the methodology that I develop in the paper can accommodate a wide range of possibilities, and, as I demonstrate below, the issue can be studied empirically.

To complete the specification of the model, I assume that the structural shocks  $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$  are independently and identically distributed (i.i.d.) Normal with covariance matrix  $\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2)$ .

The analysis in this paper assumes that the shadow rate  $r_t^*$  is only observed above the ZLB. If  $r_t^* < 0$  were observed up to scale, for instance, if we could measure QE  $b_{L,t}$  from the balance sheet of the central bank, then the computation of the likelihood would be much simpler—no filtering would be needed to deal with lags of  $r_t^* < 0$  on the right-hand side of the SVAR model introduced in the next section, but the identification problem would remain the same. More generally, we could assume that  $r_t^* < 0$  is observed with some measurement error  $\eta_t$ , and include the measurement equation  $b_{L,t} = \min(\alpha r_t^* + \eta_t, 0)$  in the model together with a specification of the distribution of  $\eta_t$ . The estimation method in this paper can then be seen as a special case where we are entirely agnostic about the measurement error. Adding such measures of unconventional policy is a potentially very useful extension of the method since, if correctly specified, they will likely improve estimation accuracy.

*Connection to the microeconometrics literature.* Equations (2) and (3) form a SEM with a limited dependent variable. The special case with  $\beta^* = 0$  in (3) can be referred to as a *kinked* SEM (KSEM), while the opposite case of  $\beta^* = \beta$  in (3) can be called a *censored* SEM (CSEM). Variants of the KSEM model have been studied in the early microeconometrics literature on limited dependent variable SEMs. Amemiya (1974) and Lee (1976) studied multivariate extensions of the well-known Tobit model (Tobin (1958)). Nelson and Olson (1978) argued that the KSEM was less suitable for microeconomic applications than the CSEM, and the latter subsequently became the main focus of the literature (Smith and Blundell (1986), Blundell and Smith (1989)). Blundell and Smith (1994) studied the unrestricted model using external instruments, so they did not consider the implications of the kink for identification.

One important lesson from the microeconometrics literature is that establishing existence and uniqueness of equilibria in this class of models is non-trivial. [Gourieroux, Laffont, and Monfort \(1980\)](#) defined a model to be “coherent” if it has a unique solution for the endogenous variables in terms of the exogenous variables, that is, if there exists a unique reduced form. More recently, the literature has distinguished between existence and uniqueness of solutions using the terms coherency and completeness of the model, respectively ([Lewbel \(2007\)](#)). Establishing coherency and completeness is a necessary first step before we can study identification and estimation.

2.1. Identification

Substituting for  $r_t^*$  in (3) using (2) and rearranging, we obtain

$$\pi_t = \tilde{c} + \tilde{\beta}(r_t - r^n) + \tilde{\varepsilon}_{1t}, \quad \text{where} \tag{6}$$

$$\tilde{\beta} = \frac{\beta - \beta^*}{1 - \gamma\beta^*}, \quad \tilde{c} = \frac{c}{1 - \gamma\beta^*} \quad \text{and} \quad \tilde{\varepsilon}_{1t} = \frac{\varepsilon_{1t} + \beta^*\varepsilon_{2t}}{1 - \gamma\beta^*}. \tag{7}$$

The system of equations (2) and (6) is now a KSEM, for which the necessary and sufficient condition for coherency and completeness (existence of a unique solution) is  $\tilde{\beta}\gamma < 1$  ([Nelson and Olson \(1978\)](#)). Using (7), the coherency and completeness condition can be expressed in terms of the structural parameters as

$$\frac{1 - \gamma\beta}{1 - \gamma\beta^*} > 0. \tag{8}$$

This condition evidently restricts the admissible range of the structural parameters. It is satisfied in the present monetary policy model, where it is natural to assume that  $\beta, \beta^* \leq 0$  and  $\gamma > 0$ . Therefore, it is possible that the coherency condition may not provide additional information relative to what is often available from natural sign restrictions on the parameters.

Under condition (8), the unique solution of the model can be written as

$$\pi_t = \mu_1 + u_{1t} - \tilde{\beta}D_t(\mu_2 + u_{2t}) \quad \text{and} \tag{9}$$

$$r_t = \max(\mu_2 + u_{2t}, 0), \tag{10}$$

where  $D_t := 1_{(r_t=0)}$  is an indicator (dummy) variable that takes the value 1 when  $r_t$  is on the boundary and zero otherwise, and

$$u_{1t} = \frac{\tilde{\varepsilon}_{1t} + \tilde{\beta}\varepsilon_{2t}}{1 - \gamma\tilde{\beta}} = \frac{\varepsilon_{1t} + \beta\varepsilon_{2t}}{1 - \gamma\beta}, \quad u_{2t} = \frac{\gamma\varepsilon_{1t} + \varepsilon_{2t}}{1 - \gamma\beta}, \tag{11}$$

$$\mu_1 = \frac{c}{1 - \gamma\beta}, \quad \mu_2 = \frac{\gamma c}{1 - \gamma\beta} + r^n.$$

Equations (9) and (10) express the endogenous variables  $\pi_t, r_t$  in terms of the exogenous variables  $\varepsilon_{1t}, \varepsilon_{2t}$ , and correspond to the decision rules of the agents in the model. It is clear that those decision rules differ in a world in which the ZLB occasionally binds, which is characterized by  $\tilde{\beta} \neq 0$ , compared to a world in which it never does (i.e., the CSEM), where  $\tilde{\beta} = 0$ . What is important for identification, however, is that in a world in which the ZLB occasionally binds, agents’ reaction to shocks differs across regimes, and

the difference depends on the parameter  $\tilde{\beta}$ , which, from equation (7), depends on the difference between the impact of conventional and unconventional policies,  $\beta$  and  $\beta^*$ , respectively. I will show that this change provides information that identifies the structural parameters: we get point identification when  $\beta^* = 0$  (the KSEM case), and partial identification when  $\beta^* \neq \beta$ . The identification argument leverages the coefficient on the kink,  $\tilde{\beta}$ , in the “incidentally kinked” regression (9), which is identified by a variant of the well-known Heckit method (Heckman (1979)). I will sketch out the argument below, and provide more details for the full SVAR model in the next section.

2.1.1. *Identification of the KSEM*

Recall that in the KSEM model,  $\tilde{\beta} = \beta$ . Consider the estimation of  $\beta$  in (4) from a regression of  $\pi_t$  on  $r_t$  using only observations above the ZLB,

$$E(\pi_t | r_t, r_t > 0) = c + \beta(r_t - r^n) + \rho \left( r_t - \mu_2 + \tau \frac{\phi(a)}{1 - \Phi(a)} \right), \quad a = \frac{-\mu_2}{\tau}, \quad (12)$$

where  $\rho = \text{cov}(u_{1t}, u_{2t})/\tau^2 - \beta = \gamma\sigma_1^2(1 - \gamma\beta)/(\gamma^2\sigma_1^2 + \sigma_2^2)$ ,  $\tau = \sqrt{\text{var}(u_{2t})}$ , and  $\phi(\cdot)$ ,  $\Phi(\cdot)$  are standard Normal density and distribution functions, respectively. The coefficient  $\rho$  is the bias in the estimation of  $\beta$  from the truncated regression (12). Now, the mean of  $\pi_t$  using the observations at the ZLB is

$$E(\pi_t | r_t = 0) = c - \beta r^n + \rho\tau \frac{\phi(a)}{\Phi(a)}. \quad (13)$$

Next, observe that  $\mu_2$ ,  $\tau$  and hence  $\phi(a)/\Phi(a)$  can be estimated from the Tobit regression (10). Therefore, we can recover the bias  $\rho$  and identify  $\beta$ . A simple way to implement this is the control function approach (Heckman (1978)). Let

$$h_t(\mu_2, \tau) := (1 - D_t)(r_t - \mu_2) - D_t \frac{\tau\phi(a)}{\Phi(a)},$$

and run the regression

$$E(\pi_t | r_t) = c_1 + \beta r_t + \rho h_t(\mu_2, \tau), \quad (14)$$

where  $c_1 = c - \beta r^n$  is an unrestricted intercept. The rank condition for the identification of  $\beta$  is simply that the regressors in (14) are not perfectly collinear. This holds if and only if  $0 < \Pr(D_t = 1) < 1$ . So, as long as some but not all the observations are at the boundary, the model is generically identified.

2.1.2. *Partial Identification of the Unrestricted SEM*

The discussion of the previous subsection shows that  $\tilde{\beta}$  is identified from the kink in the reduced-form equation for  $\pi_t$  (9). It follows from equation (7) and the order condition that  $\beta, \beta^*$  are not point-identified. I will now demonstrate that they are partially identified.

The assumption  $\text{cov}(\varepsilon_{1t}, \varepsilon_{2t}) = 0$  implies (see proof of Proposition 3 for a derivation)

$$\gamma = \frac{\omega_{12} - \omega_{22}\beta}{\omega_{11} - \omega_{12}\beta}, \quad (15)$$

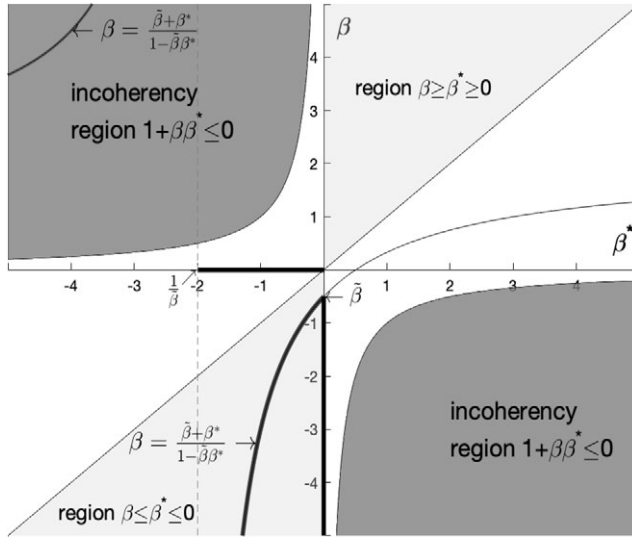


FIGURE 1.—The identified set for  $(\beta, \beta^*)$  when  $\Omega = I$  and  $\tilde{\beta} = -1/2$ , obtained by the intersection of  $\beta = \frac{\tilde{\beta} + \beta^*}{1 - \beta\beta^*}$  and  $1 + \beta\beta^* > 0$  (coherency condition). Light gray area corresponds to  $\beta \leq \beta^* \leq 0$  and  $\beta \geq \beta^* \geq 0$ . The thick part of the curve  $\beta = \frac{\tilde{\beta} + \beta^*}{1 - \beta\beta^*}$  indicates the identified set obtained from the combined restrictions, and the bold intervals on the axes give the projections of the identified set onto  $\beta$  and  $\beta^*$ .

where  $\omega_{ij} := \text{cov}(u_{it}, u_{jt})$ . Substituting for  $\gamma$  in (7) using (15) yields

$$\tilde{\beta} = \frac{\beta - \beta^*}{1 - \frac{\omega_{12} - \omega_{22}\beta}{\omega_{11} - \omega_{12}\beta}\beta^*}. \tag{16}$$

For any given value of the reduced-form parameters  $\tilde{\beta}, \Omega := \text{var}(u_t)$ ,  $u_t = (u_{1t}, u_{2t})'$ , the identified set for  $(\beta, \beta^*)$  is a one-dimensional manifold in  $\mathbb{R}^2$  defined by equation (16) intersected with the coherency condition (8).

It is instructive to illustrate the identified set graphically at some given value of  $\tilde{\beta}$  and  $\Omega$ . Consider, for example, the case  $\Omega$  equal to the identity  $I_2$ , at which (15) yields  $\gamma = -\beta$ , and the coherency condition (8) reduces to  $1 + \beta\beta^* > 0$ , and (16) yields the function  $\beta = \frac{\tilde{\beta} + \beta^*}{1 - \beta\beta^*}$ . Figure 1 plots this function at  $\tilde{\beta} = -1/2$ , and highlights in dark gray the region of incoherency defined by  $1 + \beta\beta^* \leq 0$ . The identified set is the part of the function  $\beta = \frac{\tilde{\beta} + \beta^*}{1 - \beta\beta^*}$  that lies to the right of the pole at  $1/\tilde{\beta}$ , that is, in the region  $\beta^* > 1/\tilde{\beta} = -2$  in this example.

Now, consider the additional restrictions  $\beta \geq \beta^* \geq 0$  or  $\beta \leq \beta^* \leq 0$ , highlighted by the light gray shaded areas in Figure 1. The interpretation of those restrictions is that unconventional policy neither has the opposite effect from conventional policy, nor is it more effective than conventional policy. With this additional restriction, we see that the identified set further shrinks to the part of  $\beta = \frac{\tilde{\beta} + \beta^*}{1 - \beta\beta^*}$  in the interval  $(\tilde{\beta}^{-1}, 0]$ . The projection of the identified set onto the  $\beta$  axis yields  $\beta \in (-\infty, \tilde{\beta}]$ , since  $\tilde{\beta} < 0$ , while its projection onto the  $\beta^*$  axis yields  $\beta^* \in (\tilde{\beta}^{-1}, 0]$ . Because equations (8) and (16) encapsulate all the information in the reduced-form parameters about  $\beta, \beta^*$ , the identified set obtained from them is sharp.



Let us define a new parameter  $\lambda$  such that  $\beta^* = \lambda\beta$ . The restriction indicated by the light gray areas in Figure 1 corresponds to  $\lambda \in [0, 1]$ . If we interpret  $\lambda$  as a measure of the efficacy of unconventional policy, this restriction implies that unconventional policy is neither counter- nor overproductive. This reparameterization offers a convenient way to discretize the parameter space when we compute the identified set numerically, as is the case in the more general model discussed in the next section.

In this bivariate model, it is possible to characterize the identified set analytically. Here, I discuss the identified set for  $\beta$  and defer the discussion of  $\lambda$  to Appendix A.1.

We have already established that  $\beta$  is completely unidentified when  $\tilde{\beta} = \beta^*$  (equivalently  $\lambda = 1$ ), which corresponds to the CSEM. From the definition of  $\tilde{\beta}$  in (7), it follows that  $\beta = \beta^*$  implies  $\tilde{\beta} = 0$ . So, when  $\tilde{\beta} = 0$ ,  $\beta$  is completely unidentified. It remains to see what happens when  $\tilde{\beta} \neq 0$ . Let  $\gamma_0 := \omega_{12}/\omega_{11}$ , which can be interpreted as the value the reaction function coefficient  $\gamma$  in (2) would take if  $\beta = 0$ , that is, the value corresponding to a Choleski identification scheme where  $r_t$  is placed last. In Appendix A.1, I prove the following bounds:

$$\begin{aligned}
 &\text{if } \tilde{\beta} = 0 \text{ or } \tilde{\beta}\gamma_0 < 0, \text{ then } \beta \in \mathfrak{R}; \text{ otherwise} \\
 &\text{if } \omega_{12} = \gamma_0 = 0, \text{ then } \beta \in (-\infty, \tilde{\beta}] \text{ if } \tilde{\beta} < 0 \text{ or } \beta \in [\tilde{\beta}, \infty) \text{ if } \tilde{\beta} > 0; \\
 &\text{if } 0 < \tilde{\beta}\gamma_0 \leq 1, \text{ then } \beta \in \left[\frac{1}{\gamma_0}, \tilde{\beta}\right] \text{ if } \tilde{\beta} < 0 \text{ or } \beta \in \left[\tilde{\beta}, \frac{1}{\gamma_0}\right] \text{ if } \tilde{\beta} > 0; \\
 &\text{if } \tilde{\beta}\gamma_0 > 1, \text{ then } \lambda < 0.
 \end{aligned} \tag{17}$$

We see that when  $\tilde{\beta} \neq 0$  and  $0 \leq \tilde{\beta}\gamma_0 \leq 1$ , we can identify both the sign of the causal effect  $\beta$  of  $r_t$  on  $\pi_t$  and get bounds on its magnitude. In particular, the identified coefficient  $\tilde{\beta}$  is an attenuated measure of the true causal effect  $\beta$ . Moreover,  $\tilde{\beta}\gamma_0 > 1$  implies that  $\beta^*$  has the opposite sign from  $\beta$ , that is, unconventional policy has the opposite effect of the conventional one. That could be interpreted as saying that unconventional policy is counterproductive.

Finally, the hypothesis that unconventional policy is as effective as conventional policy,  $\beta^* = \beta$  or  $\lambda = 1$ , is equivalent to the null hypothesis  $H_0 : \tilde{\beta} = 0$ . The alternative that unconventional policy is less effective than conventional policy,  $\beta > \beta^*$  if  $\beta > 0$ , or  $\beta < \beta^*$  if  $\beta < 0$ , corresponds to the two-sided alternative  $H_1 : \beta \neq \beta^*$ . This can be tested using a likelihood ratio test.

### 3. SVAR WITH AN OCCASIONALLY BINDING CONSTRAINT

I now develop the methodology for identification and estimation of SVARs with an occasionally binding constraint. Let  $Y_t = (Y'_{1t}, Y_{2t})'$  be a vector of  $k$  endogenous variables, partitioned such that the first  $k - 1$  variables  $Y_{1t}$  are unrestricted and the  $k$ th variable  $Y_{2t}$  is bounded from below by  $b$ .<sup>4</sup> Define the latent process  $Y_{2t}^*$  that is only observed, and equal to  $Y_{2t}$ , whenever  $Y_{2t} > b$ . If  $Y_{2t}$  is a policy instrument,  $Y_{2t}^*$  can be thought of as the “shadow” instrument that measures the desired policy stance. The  $p$ th-order SVAR

<sup>4</sup>The lower bound does not need to be constant. All we need is to observe the periods in which the economy is at the ZLB regime.

model is given by the equations

$$A_{11}Y_{1t} + A_{12}Y_{2t} + A_{12}^*Y_{2t}^* = B_{10}X_{0t} + \sum_{j=1}^p B_{1,j}Y_{t-j} + \sum_{j=1}^p B_{1,j}^*Y_{2,t-j}^* + \varepsilon_{1t}, \tag{18}$$

$$A_{22}^*Y_{2t}^* + A_{22}Y_{2t} + A_{21}Y_{1t} = B_{20}X_{0t} + \sum_{j=1}^p B_{2,j}Y_{t-j} + \sum_{j=1}^p B_{2,j}^*Y_{2,t-j}^* + \varepsilon_{2t}, \tag{19}$$

$$Y_{2t} = \max(Y_{2t}^*, b),$$

for  $t \geq 1$  given a set of initial values  $Y_{-s}, Y_{2,-s}^*$ , for  $s = 0, \dots, p - 1$ , and  $X_{0t}$  are exogenous and predetermined variables.

Equation (19) can be interpreted as a policy reaction function because it determines the desired policy stance  $Y_{2t}^*$ . Similarly,  $\varepsilon_{2t}$  is the corresponding policy shock. The above model is a dynamic SEM. Two important differences from a standard SEM are the presence of (i) latent lags amongst the predetermined variables on the right-hand side, which complicates estimation; and (ii) the contemporaneous value of  $Y_{2t}$  in the policy reaction function (19), which allows it to vary across ZLB and non-ZLB regimes. The presence of latent lags  $Y_{2,t-j}^*$  in the policy rule (19) is particularly useful because it allows the model to incorporate forward guidance (Reifschneider and Williams (2000), Debortoli, Gali, and Gambetti (2019)); see the Supplemental Material for details.

Collecting all the observed predetermined variables  $X_{0t}, Y_{t-1}, \dots, Y_{t-p}$  into a vector  $X_t$ , and the latent lags  $Y_{2,t-1}^*, \dots, Y_{2,t-p}^*$  into  $X_t^*$ , and similarly for their coefficients, the model can be written compactly as

$$\begin{pmatrix} A_{11} & A_{12}^* & A_{12} \\ A_{21} & A_{22}^* & A_{22} \end{pmatrix} \begin{pmatrix} Y_{1t} \\ Y_{2t}^* \\ Y_{2t} \end{pmatrix} = BX_t + B^*X_t^* + \varepsilon_t, \tag{20}$$

$$Y_{2t} = \max\{Y_{2t}^*, b\}.$$

The vector of structural errors  $\varepsilon_t$  is assumed to be i.i.d. Normally distributed with zero mean and identity covariance.

In the previous section, we defined the KSEM as a special case of the general model, where  $Y_{2t}^* < b$  has no (contemporaneous) impact on  $Y_{1t}$ . In the dynamic setting, it feels natural to define the corresponding “kinked SVAR” model (KSVAR) as a model in which  $Y_{2t}^*$  has neither contemporaneous nor dynamic effects. Therefore, the KSVAR obtains as a special case of (20) when both  $A_{12}^* = 0$ , and  $B^* = 0$ , which corresponds to a situation in which the bound is fully effective in constraining what policy can achieve at all horizons.

The opposite extreme to the KSVAR is the censored SVAR model (CSVAR). Again, unlike the CSEM, which only characterizes contemporaneous effects, the idea of a CSVAR is to impose the assumption that the constraint is irrelevant at all horizons. So, it corresponds to a fully unrestricted linear SVAR in the latent process  $(Y_{1t}', Y_{2t}^*)'$ . This is a special case of (20) when both  $A_{12} = 0$  and the elements of  $B$  corresponding to lagged  $Y_{2t}$  are equal to zero. Finally, I refer to the general model given by (20) as the “censored and kinked SVAR” (CKSVAR).

Define the  $k \times k$  square matrices

$$\bar{A} := \begin{pmatrix} A_{11} & A_{12} + A_{12}^* \\ A_{21} & A_{22} + A_{22}^* \end{pmatrix} \quad \text{and} \quad A^* := \begin{pmatrix} A_{11} & A_{12}^* \\ A_{21} & A_{22}^* \end{pmatrix}. \tag{21}$$

$\overline{A}$  determines the impact effects of structural shocks during periods when the constraint does not bind.  $A^*$  does the same for periods when the constraint binds.

To analyze the CKSVAR, we first need to establish existence and uniqueness of the reduced form. This is done in the following proposition.

PROPOSITION 1: *The model given in equation (20) is coherent and complete (i.e., it has a unique solution) if and only if*

$$\kappa := \frac{\overline{A}_{22} - A_{21}A_{11}^{-1}\overline{A}_{12}}{A_{22}^* - A_{21}A_{11}^{-1}A_{12}^*} > 0. \tag{22}$$

Note that (22) does not depend on the coefficients on the lags (whether latent or observed), so it is exactly the same as in a static SEM. This condition is useful for inference, for example, when constructing confidence intervals or posteriors, because it restricts the range of admissible values for the structural parameters. It can also be checked empirically when the structural parameters are point-identified.

If condition (22) is satisfied, there exists a reduced-form representation of the CKSVAR model (20). For convenience of notation, define the indicator (dummy variable) that takes the value 1 if the constraint binds and zero otherwise:

$$D_t = 1_{\{Y_{2t}=b\}}. \tag{23}$$

PROPOSITION 2: *If (22) holds, and for any initial values  $Y_{-s}, Y_{2,-s}^*, s = 0, \dots, p - 1$ , the reduced-form representation of (20) for  $t \geq 1$  is given by*

$$Y_{1t} = \overline{C}_1 X_t + \overline{C}_1^* \overline{X}_t^* + u_{1t} - \tilde{\beta} D_t (\overline{C}_2 X_t + \overline{C}_2^* \overline{X}_t^* + u_{2t} - b), \tag{24}$$

$$Y_{2t} = \max(\overline{Y}_{2t}^*, b), \tag{25}$$

$$\overline{Y}_{2t}^* = \overline{C}_2 X_t + \overline{C}_2^* \overline{X}_t^* + u_{2t}, \tag{26}$$

$$Y_{2t}^* = (1 - D_t) \overline{Y}_{2t}^* + D_t (\kappa \overline{Y}_{2t}^* + (1 - \kappa)b), \tag{27}$$

where  $u_t = (u'_{1t}, u'_{2t})' = \overline{A}^{-1} \varepsilon_t$ ,  $\overline{C}^* = (\overline{C}_1^*, \overline{C}_2^*)' = \kappa \overline{A}^{-1} B^*$ ,  $\overline{X}_t^* = (\overline{x}_{t-1}, \dots, \overline{x}_{t-p})'$ ,  $\overline{x}_t = \min(\overline{Y}_{2t}^* - b, 0)$ ,  $\overline{x}_{-s} = \kappa^{-1} \min(Y_{2,-s}^* - b, 0)$ ,  $s = 0, \dots, p - 1$ ,

$$\tilde{\beta} = (A_{11} - A_{12}^* A_{22}^{*-1} A_{21})^{-1} (A_{12}^* A_{22}^{*-1} A_{22} - A_{12}), \tag{28}$$

$\kappa$  is defined in (22), and the matrices  $\overline{C}_1, \overline{C}_2$ , are given in equation (49) in the *Appendix*.

Note that the “reduced-form; latent process  $\overline{Y}_{2t}^*$  is, in general, different from the “structural” shadow rate  $Y_{2t}^*$  defined by (27). They coincide only when  $\kappa = 1$ . This holds, for example, in the CSVAR model.

Equation (25) combined with (26) is a familiar dynamic Tobit regression model with the added complexity of latent lags included as regressors whenever  $\overline{C}_2^* \neq 0$ . Likelihood estimation of the univariate version of this model was studied by Lee (1999). The  $k - 1$  equations (24) are “incidentally kinked” dynamic regressions, that I have not seen analyzed before.

### 3.1. Identification

#### 3.1.1. Identification of Reduced-Form Parameters

Let  $\psi$  denote the parameters that characterize the reduced form (24)–(25):  $\tilde{\beta}, \bar{C}, \bar{C}^*$ , and  $\Omega = \text{var}(u_t)$ . It is useful to decompose  $\psi$  into  $\psi_2 = (\bar{C}_2, \bar{C}_2^*, \tau)'$ , where  $\tau = \sqrt{\text{var}(u_{2t})}$ , and  $\psi_1 = (\text{vec}(\bar{C}_1)', \text{vec}(\bar{C}_1^*)', \tilde{\beta}', \delta', \text{vech}(\Omega_{1,2}))'$ , where  $\delta = \Omega_{12}/\tau^2$ ,  $\Omega_{1,2} = \Omega_{11} - \delta\delta'\tau^2$ , and  $\Omega_{ij} = \text{cov}(u_{it}, u_{jt})$ .

Equation (25) is the dynamic Tobit regression model studied by Lee (1999). So, its parameters,  $\psi_2$ , are generically identified provided that the regressors are not perfectly collinear. This requires that  $0 < \Pr(D_t = 1) < 1$ .

Given  $\psi_2$ , the identification of the remaining parameters,  $\psi_1$ , can be characterized using a control function approach. Consider the  $k - 1$  regression equations

$$E(Y_{1t}|Y_{2t}, X_t, \bar{X}_t^*) = \bar{C}_1 X_t + \bar{C}_1^* \bar{X}_t^* + \tilde{\beta} Z_{1t} + \delta Z_{2t}, \tag{29}$$

where

$$Z_{1t} = D_t \left( b - \bar{C}_2 X_t - \bar{C}_2^* \bar{X}_t^* - \frac{\tau \phi(a_t)}{\Phi(a_t)} \right), \tag{30}$$

$$Z_{2t} = (1 - D_t)(Y_{2t} - \bar{C}_2 X_t - \bar{C}_2^* \bar{X}_t^*) + D_t \frac{\tau \phi(a_t)}{\Phi(a_t)}, \tag{31}$$

$a_t = \left( \frac{b - \bar{C}_2 X_t - \bar{C}_2^* \bar{X}_t^*}{\tau} \right)$ , and  $\phi(\cdot), \Phi(\cdot)$  are the standard Normal density and distribution functions, respectively. When  $\bar{C}^*$  is different from zero, regressors  $\bar{X}_t^*, Z_{1t}$ , and  $Z_{2t}$  in (29) are unobserved, so we need to replace them with their expectations conditional on  $Y_{2t}, Y_{t-1}, \dots, Y_1$ . Then, the regressors on the right-hand side of (29) become  $\mathbf{X}_t := (X_t', \bar{X}_{t|t}^*, Z_{1t|t}, Z_{2t|t})'$ , where  $h_{t|t} := E(h(\bar{X}_t^*)|Y_{2t}, Y_{t-1}, \dots, Y_1)$  for any function  $h(\cdot)$  whose expectation exists.<sup>5</sup> The coefficients  $\bar{C}_1, \bar{C}_1^*, \tilde{\beta}$ , and  $\delta$  are generically identified if the regressors  $\mathbf{X}_t$  are not perfectly collinear.

#### 3.1.2. Identification of Structural Parameters

From the order condition, we can easily establish that there are not enough restrictions to identify all the structural parameters in the CKSVAR (20). Let  $k_0 = \dim(X_{0t})$  denote the number of predetermined variables other than the own lags of  $Y_t$ . For example, in a standard VAR without deterministic trends, we have  $X_{0t} = 1$ , so  $k_0 = 1$ . The number of reduced-form parameters  $\psi$  is  $k_0 k + k^2 p$  (in  $\bar{C}$ ) plus  $k p$  (in  $\bar{C}^*$ ) plus  $k - 1$  (in  $\tilde{\beta}$ ) plus  $k(k + 1)/2$  (in  $\Omega$ ). The number of structural parameters in (20) is  $k_0 k + k^2 p$  (in  $B$ ) plus  $k p$  (in  $B^*$ ) plus  $k^2$  (in  $\bar{A}$ ) plus  $k$  (in  $A_{12}^*$  and  $A_{22}$ ). So, the CKSVAR is underidentified by  $k(k - 1)/2 + 1$  restrictions. Nevertheless, I will show that the impulse responses to  $\varepsilon_{2t}$  are identified. Specifically, they are point-identified when  $A_{12}^* = 0$ , and partially identified when  $A_{12}^* \neq 0$  but  $A_{12}^*$  and  $A_{12}$  have the same sign, analogous to the bounds given in equation (17) in the previous section.

Because the CKSVAR is nonlinear, IRFs are obviously state-dependent, and there are many ways one can define them; see Koop, Pesaran, and Potter (1996). The IRF to  $\varepsilon_{2t}$ ,

<sup>5</sup>In the KSVAR model, we have  $\bar{C}_1^* = 0$  and  $\bar{C}_2^* = 0$ , so  $\bar{X}_t^*$  drops out of (29), and the regressors  $Z_{1t}, Z_{2t}$  are observed, so  $Z_{j|t} = Z_{jt}, j = 1, 2$ .

according to any of the definitions proposed in the literature, is identified if the reduced-form errors  $u_t$  can be expressed as a known function of  $\varepsilon_{2t}$  and a process that is orthogonal to it, that is,  $u_t = g(\varepsilon_{2t}, e_t)$ , where  $e_t$  is independent of  $\varepsilon_{2t}$ . From Proposition 2, it follows that the function  $g$  is linear, and more specifically,

$$u_{1t} = (I_{k-1} - \bar{\beta}\bar{\gamma})^{-1}(\bar{\varepsilon}_{1t} + \bar{\beta}\bar{\varepsilon}_{2t}) \quad \text{and} \tag{32}$$

$$u_{2t} = (1 - \bar{\gamma}\bar{\beta})^{-1}(\bar{\varepsilon}_{2t} + \bar{\gamma}\bar{\varepsilon}_{1t}), \tag{33}$$

where

$$\begin{aligned} \bar{\beta} &:= -A_{11}^{-1}\bar{A}_{12}, & \bar{\gamma} &:= -\bar{A}_{22}^{-1}A_{21}, \\ \bar{\varepsilon}_{1t} &:= A_{11}^{-1}\varepsilon_{1t}, & \bar{\varepsilon}_{2t} &:= \bar{A}_{22}^{-1}\varepsilon_{2t}, \end{aligned} \tag{34}$$

and  $\bar{A}_{22} = A_{22}^* + A_{22}$ , defined in (21). Note that  $\bar{\beta}$  can be interpreted as the response of  $Y_{1t}$  to a shock that increases  $Y_{2t}$  by one unit, and  $\bar{\gamma}$  are the contemporaneous reaction function coefficients of  $Y_{2t}$  to  $Y_{1t}$  when  $Y_{2t} > b$  (unconstrained regime). The shock vector  $\bar{\varepsilon}_{1t}$  is not structural but it is orthogonal to  $\varepsilon_{2t}$ , so it plays the role of  $e_t$  in  $u_t = g(\varepsilon_{2t}, e_t)$ . Hence, the IRF is identified if and only if  $\bar{\beta}$ ,  $\bar{\gamma}$ , and  $\bar{A}_{22}$  are identified.

The following proposition shows identification when  $A_{12}^* = 0$ .

PROPOSITION 3: *When  $A_{12}^* = 0$  and the coherency condition (22) holds, the parameters in (32)–(33) are identified by the equations  $\bar{\beta} = \tilde{\beta}$ ,*

$$\bar{\gamma} = (\Omega'_{12} - \Omega_{22}\bar{\beta}')(\Omega_{11} - \Omega_{12}\bar{\beta}')^{-1} \quad \text{and} \tag{35}$$

$$\bar{A}_{22}^{-1} = \sqrt{(-\bar{\gamma}, 1)\Omega(-\bar{\gamma}, 1)'} \tag{36}$$

REMARKS: 1.  $\bar{\beta} = \tilde{\beta}$  follows immediately from the definition (28) with  $A_{12}^* = 0$ . Equations (35) and (36) hold without the restriction  $A_{12}^* = 0$ . They follow from the orthogonality of the shocks  $\varepsilon_{2t}$  and  $\bar{\varepsilon}_{1t}$ .

2. An instrumental variables interpretation of this identification result is as follows. Define the instrument

$$Z_t := Y_{1t} - \tilde{\beta}Y_{2t} = A_{11}^{-1}B_1X_t + A_{11}^{-1}B_1^*X_t^* + A_{11}^{-1}\varepsilon_{1t},$$

where the second equality holds when  $A_{12}^* = 0$ . The orthogonality of the errors  $E(\varepsilon_{1t}\varepsilon_{2t}) = 0$  implies  $E(Z_t\varepsilon_{2t}) = 0$ . So,  $Z_t$  are valid  $k - 1$  instruments for the  $k - 1$  endogenous regressors  $Y_{1t}$  in the structural equation of  $Y_{2t} = \max(Y_{2t}^*, b)$ , where  $Y_{2t}^*$  is given by (19). Normalizing (19) in terms of  $Y_{2t}^*$  yields the structural equation in the more familiar form of a policy rule:

$$Y_{2t} = \max(\bar{\gamma}Y_{1t} + \bar{B}_2X_t + \bar{B}_2^*X_t^* + \bar{\varepsilon}_{2t}, b), \tag{37}$$

where  $\bar{B}_2 = \bar{A}_{22}^{-1}B_2$ ,  $\bar{B}_2^* = \bar{A}_{22}^{-1}B_2^*$ . Since  $A_{11}^{-1}$  is nonsingular, the coefficient matrix of  $Z_t$  in the “first-stage” regressions of  $Y_{1t}$  is nonsingular, so the coefficients of (37) are generically identified by the rank condition. An alternative to the Tobit IV regression model (37) is

the indirect Tobit regression approach used in the static SEM by [Blundell and Smith \(1994\)](#). Equation (37) can be written as the dynamic Tobit regression

$$Y_{2t} = \max(\tilde{\gamma}Z_t + \tilde{B}_2X_t + \tilde{B}_2^*X_t^* + \tilde{\varepsilon}_{2t}, b), \tag{38}$$

where  $\tilde{\gamma} = (1 - \bar{\gamma}\bar{\beta})^{-1}\bar{\gamma}$ ,  $\tilde{B}_2 = (1 - \bar{\gamma}\bar{\beta})^{-1}\bar{B}_2$ ,  $\tilde{B}_2^* = (1 - \bar{\gamma}\bar{\beta})^{-1}\bar{B}_2^*$ , and  $\tilde{\varepsilon}_{2t} = (1 - \bar{\gamma}\bar{\beta})^{-1}\bar{\varepsilon}_{2t}$ . Note that the coherency condition (22) becomes  $\kappa = \frac{A_{22}}{A_{22}^*}(1 - \bar{\gamma}\bar{\beta}) > 0$ , so  $1 - \bar{\gamma}\bar{\beta} \neq 0$ , which guarantees the existence of the representation (38). Given  $\bar{\beta} = \tilde{\beta}$ , the structural parameter  $\bar{\gamma}$  can then be obtained as  $\bar{\gamma} = \tilde{\gamma}(I_{k-1} + \tilde{\beta}\tilde{\gamma})^{-1}$ , and similarly for the remaining structural parameters in (37).

3. The parameter  $A_{22}$  allows the reaction function of  $Y_{2t}^*$  to differ across the two regimes. The special case  $A_{22} = 0$  thus corresponds to the restriction that the reaction function remains the same across regimes. The parameters  $A_{22}$  and  $A_{22}^*$  are not separately identified. Hence,  $A_{22}^{*-1}$ , the scale of the response to the shock  $\varepsilon_{2t}$  during periods when  $Y_{2t} = b$ , is not identified.<sup>6</sup> Similarly,  $\kappa = \frac{A_{22}}{A_{22}^*}(1 - \bar{\gamma}\bar{\beta})$  is not identified, and therefore, neither is the structural shadow value  $Y_{2t}^*$  in equation (27). Identification of these requires an additional restriction on  $A_{22}$ , for example,  $A_{22} = 0$ . Turning this discussion around, we see that a change in the reaction function across regimes does not destroy the point identification of the effects of policy during the unconstrained regime, since the latter only requires  $\bar{\beta}$ ,  $\bar{\gamma}$ , and  $\bar{A}_{22}$ , not  $A_{22}^*$  or  $\kappa$ .

Next, we turn to the case  $A_{12}^* \neq 0$ , and derive identification under restrictions on the sign and magnitude of  $A_{12}^*$  relative to  $A_{12}$  and  $A_{12}^*$  relative to  $A_{22}$ . The first restriction is motivated by a generalization of the discussion on the SEM model in Section 2.1.2. Specifically, if  $\bar{A}_{12} = A_{12} + A_{12}^*$  measures the effect of conventional policy (operating in the unconstrained regime) and  $A_{12}^*$  measures the effect of unconventional policy (operating in the constrained regime), then the assumption that  $A_{12}$  and  $A_{12}^*$  have the same sign means that unconventional policy effects are neither in the opposite direction nor larger in absolute value than conventional policy effects. In other words, unconventional policy is neither counterproductive nor overproductive relative to conventional policy. This can be characterized by the specification  $A_{12}^* = \Lambda\bar{A}_{12}$  and  $A_{12} = (I_{k-1} - \Lambda)\bar{A}_{12}$ , where  $\Lambda = \text{diag}(\lambda_j)$ ,  $\lambda_j \in [0, 1]$  for  $j = 1, \dots, k - 1$ . I further impose the restriction that  $\lambda_j = \lambda$  for all  $j$ , so that  $A_{12}^* = \lambda\bar{A}_{12}$  and  $A_{12} = (1 - \lambda)\bar{A}_{12}$  with  $\lambda \in [0, 1]$ . This, in turn, means that  $Y_{2t}$  and  $Y_{2t}^*$  enter each of the first  $k - 1$  structural equations for  $Y_{1t}$  only via the common linear combination  $\lambda Y_{2t}^* + (1 - \lambda)Y_{2t}$ , which can be interpreted as a measure of the effective policy stance.

We also need to consider the impact of  $A_{22}$  on identification. The parameter  $\zeta = \bar{A}_{22}/A_{22}^*$  gives the ratio of the standard deviation of the monetary policy shock in the constrained relative to the unconstrained regime. It is also the ratio of the reaction function coefficients in the two regimes, for example,  $A_{22}^{*-1}A_{21}$  versus  $\bar{A}_{22}^{-1}A_{21}$ . I will impose  $\zeta > 0$ , so that the sign of the policy shock does not change across regimes. With the above reparameterization and the definitions in (34), the identified coefficient  $\tilde{\beta}$  in (7) can be written as

$$\tilde{\beta} = (1 - \xi)(I - \xi\bar{\beta}\tilde{\gamma})^{-1}\bar{\beta}, \quad \xi := \lambda\zeta. \tag{39}$$

<sup>6</sup>This is akin to the well-known property of a probit model that the scale of the distribution of the latent process is not identifiable.

Similarly, given  $\zeta > 0$ , the coherency condition (22) reduces to  $(1 - \bar{\gamma}\bar{\beta})(1 - \xi\bar{\gamma}\bar{\beta}) > 0$ . Notice that the parameters  $\lambda, \zeta$  only appear multiplicatively, so it suffices to consider them together as  $\xi = \lambda\zeta$ . Once  $\bar{\beta}$  is known, the remaining structural parameters needed to obtain the IRF to  $\varepsilon_{2t}$  are  $\bar{\gamma}$  and  $\bar{A}_{22}$ , and they are obtained from Proposition 3. So, the identified set can be characterized by varying  $\xi$  over its admissible range. Without further restrictions on  $\zeta$ , the admissible range is obviously  $\xi \geq 0$ . If we further assume that  $\zeta \leq 1$ , that is, that the slope of the reaction function coefficients is no steeper in the constrained regime than in the unconstrained regime, then  $\xi \in [0, 1]$ , and so partial identification proceeds exactly along the lines of the SEM in the previous section where  $\lambda$  played the role of  $\xi$ . In the case  $k = 2$ , the bounds derived in equation (17) apply, with  $\beta = \bar{\beta}$  in the notation of the present section. However, when  $k > 2$ , it is difficult to obtain a simple analytical characterization of the identified set for  $\bar{\beta}$ . In any case, we will typically wish to obtain the identified set for functions of the structural parameters, such as the IRF. This can be done numerically by searching over a fine discretization of the admissible range for  $\xi$ . An algorithm for doing this is provided in Appendix E.2 of the Supplemental Material.

### 3.2. Estimation

Estimation of the CKSVAR is carried out by Maximum Likelihood (ML) using either a version of the sequential importance sampler (SIS) of Lee (1999) or the fully adapted particle filter (FAPF) of Malik and Pitt (2011) to evaluate the likelihood, except in the case of the KSVAR model for which the likelihood is available analytically. The details are given in Appendix E.1 of the Supplemental Material.

Using the limit theory of Newey and McFadden (1994), the ML estimator can be shown to be consistent and asymptotically Normal and the LR statistic asymptotically  $\chi^2$  with degrees of freedom equal to the number of restrictions. Standard asymptotics arise when the probability of each regime occurring is bounded away from zero. Infrequent visits to one of the two regimes will slow down the rate of convergence of the estimator, but will not lead to a non-standard limiting distribution. Since the focus of this paper is on identification, I will not discuss primitive conditions for these results, such as geometric ergodicity, which can be shown, for example, by bounding the joint spectral radius of the companion-form representation of the model (Liescher (2005)). Instead, I report Monte Carlo simulation results on the finite-sample properties of ML estimators and LR tests in Appendix B. They show that the Normal distribution provides a very good approximation to the finite-sample distribution of the ML estimators. I find some finite-sample size distortion in the LR tests of various restrictions on the CKSVAR, but this can be addressed effectively with a parametric bootstrap, as shown in the Appendix.

One interesting observation from the simulations is that the LR test of the CSVAR restrictions against the CKSVAR appears to be less powerful than the corresponding test of the KSVAR restrictions against the CKSVAR. Thus, we expect to be able to detect deviations from KSVAR more easily than deviations from CSVAR. In other words, finding evidence against the hypothesis that unconventional policies are fully effective (CSVAR) may be harder than finding evidence against the opposite hypothesis that they are completely ineffective (KSVAR).

## 4. APPLICATION

I use the three-equation SVAR of Stock and Watson (2001), consisting of inflation, the unemployment rate, and the Federal Funds rate to provide a simple empirical illustration

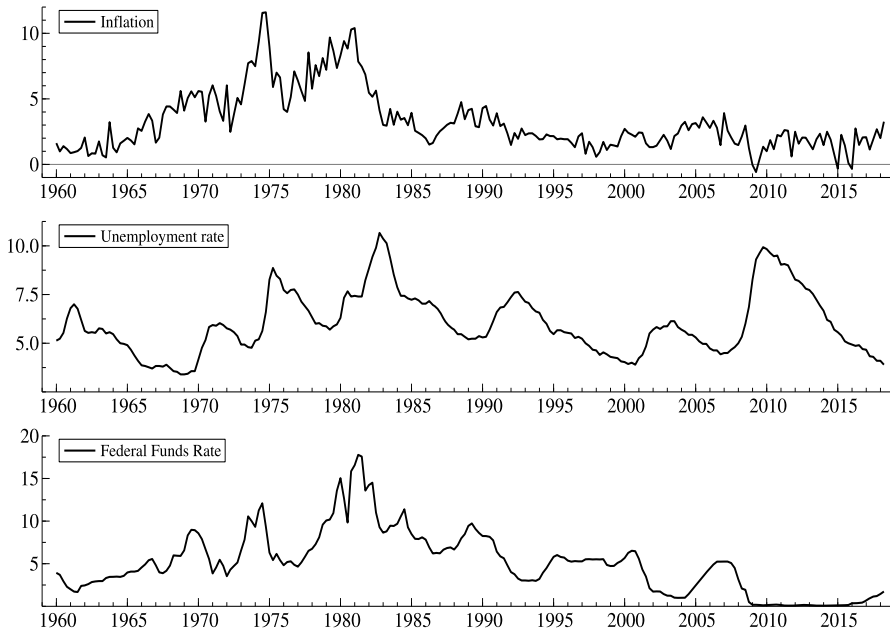


FIGURE 2.—Data used in Stock and Watson (2001) over the extended sample 1960q1 to 2018q2.

of the methodology developed in this paper. As discussed in Stock and Watson (2001), this model is far too limited to provide credible identification of structural shocks, so the results in this section are meant as an illustration of the new methods.

The data are quarterly and are constructed exactly as in Stock and Watson (2001).<sup>7</sup> The variables are plotted in Figure 2 over the extended sample 1960q1 to 2018q2. I will consider all periods in which the Fed Funds rate was below 20 basis points to be on the ZLB. This includes 28 quarters, or 11% of the sample.

#### 4.1. Tests of Efficacy of Unconventional Policy

I estimate three specifications of the SVAR(4) with the ZLB: the unrestricted CKSVAR specification, as well as the restricted KSVAR and CSVAR specifications. The maximum log-likelihood for each model is reported in Table I, computed using the SIS algorithm in the case of CKSVAR and CSVAR, with 1000 particles. The accuracy of the SIS algorithm was gauged by comparing the log-likelihood to the one obtained using the resampling FAPF algorithm. In both CKSVAR and CSVAR, the difference is very small. The results are also very similar when we increase the number of particles to 10,000. Finally, the table reports the LR tests of KSVAR and CSVAR against CKSVAR using both asymptotic and parametric bootstrap  $p$ -values.

The KSVAR imposes the restriction that no latent lags (i.e., lags of the shadow rate) should appear on the right-hand side of the model, that is,  $B^* = 0$  in (20) or  $\bar{C}_1^* = 0$  and  $\bar{C}_2^* = 0$  in (24) and (25). This amounts to 12 exclusion restrictions on the CKSVAR(4),

<sup>7</sup>The inflation data are computed as  $\pi_t = 400 \ln(P_t/P_{t-1})$ , where  $P_t$  is the implicit GDP deflator and  $u_t$  is the civilian unemployment rate. Quarterly data on  $u_t$  and  $i_t$  are formed by taking quarterly averages of their monthly values.



TABLE I  
ESTIMATED CKSVAR MODELS IN INFLATION, UNEMPLOYMENT, AND THE FEDERAL FUNDS RATE.<sup>a</sup>

Model	log lik	(FAPF)	# Restr.	LR Stat.	Asym. $p$ -val.	Boot. $p$ -val.
CKSVAR(4)	-81.64	-81.94				
KSVAR(4)	-97.05	-	12	30.82	0.002	0.011
CSVAR(4)	-94.86	-94.87	14	26.43	0.023	0.117

<sup>a</sup>Maximized log-likelihood of various SVAR models in inflation, unemployment, and Federal Funds rate. CKSVAR corresponds to the unrestricted specification (24)–(25); KSVAR excludes latent lags; CSVAR is a purely censored model. CKSVAR and CSVAR likelihoods computed using sequential importance sampling with 1000 particles (alternative estimates based on Fully Adapted Particle Filtering with resampling are shown in parentheses). Asymptotic  $p$ -values from  $\chi^2_q$ ,  $q$  = number of restrictions. Bootstrap  $p$ -values from parametric bootstrap with 999 replications. Sample: 1960q1-2018q2 (234 obs, 11% at ZLB).

four restrictions in each of the three equations. This is necessary (but not sufficient) for the hypothesis that unconventional policy is completely ineffective at all horizons. It is necessary because  $\bar{C}^* = (\bar{C}_1^*, \bar{C}_2^*)' \neq 0$  would imply that unconventional policy has at least a lagged effect on  $Y_t$ .  $\bar{C}^* = 0$  is not sufficient to infer that unconventional policy is completely ineffective because it may still have a contemporaneous effect on  $Y_{1t}$  if  $A_{12}^* \neq 0$ , and the latter is not point-identified. The result of the test in Table I shows that lags of the shadow rate are statistically significant at the 5% level, rejecting the null hypothesis that unconventional policy has no effect.

The CSVAR model imposes the restriction that only the coefficients on the lags of the shadow rate (which is equal to the actual rate above the ZLB) are different from zero in the model, that is, the elements of  $B$  corresponding to lags of  $Y_{2t}$  in (20) are all zero, or equivalently, the elements of  $\bar{C}$  corresponding to lags of  $Y_{2t}$  in (24) and (25) are all equal to  $\bar{C}^*$ . In addition, it imposes the restriction that  $\tilde{\beta} = 0$  in (24), that is, no kink in the reduced-form equations for inflation and unemployment across regimes, yielding 14 restrictions in total. This is necessary for the hypothesis that the ZLB is empirically irrelevant for policy in that it does not limit what monetary policy can achieve. The evidence against this hypothesis is not as strong as in the case of the KSVAR. The asymptotic  $p$ -value is 0.023, indicating rejection of the null hypothesis that unconventional policy is as effective as conventional policy at the 5% level, but the bootstrap  $p$ -value is 0.117. Note that this difference could also be due to the fact that the test of the CSVAR restrictions may be less powerful than the test of the KSVAR restrictions, as indicated by the simulations reported in the previous section. Thus, I would cautiously conclude that the evidence on the empirical relevance of the ZLB is mixed. Further evidence on the efficacy of unconventional policy will also be provided in the next subsection.

#### 4.2. Impulse Response Functions

Based on the evidence reported in the previous section, I estimate the IRFs associated with the monetary policy shock using the unrestricted CKSVAR specification, and compare them to recursive IRFs from the CSVAR specification that place the Federal Funds rate last in the causal ordering. From the identification results in Section 3, the CKSVAR point-identifies the nonrecursive IRFs only under the assumption that the shadow rate has no contemporaneous effect of  $Y_{1t}$ , that is,  $A_{12}^* = 0$  in (18). Note that, due to the non-linearity of the model, the IRFs are state-dependent. I use the following definition of the

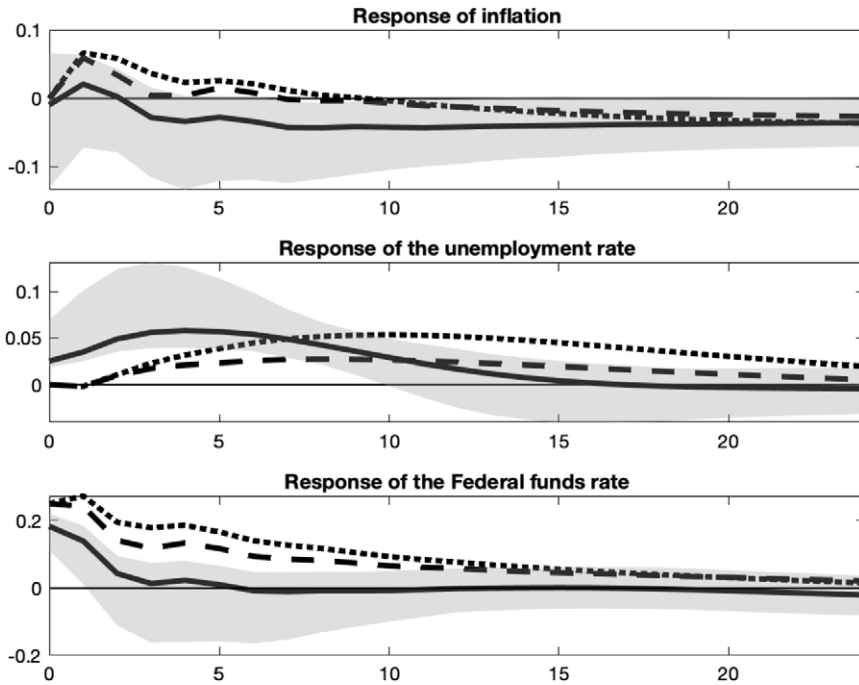


FIGURE 3.—IRFs to a 25bp monetary policy shock in 2018q3 from a CKSVAR(4) estimated over the period 1960q1 to 2018q2. The solid line corresponds to point estimates from the nonrecursive identification using the ZLB under the assumption that unconventional policy is ineffective on impact, with 90% bootstrap error bands in gray. The dashed line corresponds to the nonlinear recursive IRF, estimated with the CSVAR(4) under the restriction that the contemporaneous impact of Fed Funds on inflation and unemployment is zero. The dotted line corresponds to the recursive IRF from a linear SVAR(4) estimated by OLS with Fed Funds ordered last.

IRF from [Koop, Pesaran, and Potter \(1996\)](#):<sup>8</sup>

$$IRF_{h,t}(s, X_t, \bar{X}_t^*) = E(Y_{t+h} | \varepsilon_{2t} = s, X_t, \bar{X}_t^*) - E(Y_{t+h} | \varepsilon_{2t} = 0, X_t, \bar{X}_t^*). \quad (40)$$

Figure 3 reports the nonrecursive IRFs to a 25 basis points monetary policy shock at the end of the sample, 2018q3 (at which point  $\bar{X}_t^* = 0$  in (40) because interest rates had been above the ZLB over the previous four quarters), from the CKSVAR under the assumption that unconventional policy has no contemporaneous effect ( $\lambda = 0$ ). It also reports two different estimates of recursive IRFs using the identification scheme in [Stock and Watson \(2001\)](#) with interest rates placed last. The first estimate is obtained from the CSVAR specification, and the second is a “naive” OLS estimate of the IRF that ignores the ZLB constraint—a direct application of the method in [Stock and Watson \(2001\)](#) to the present sample. The figure also reports 90% bootstrap error bands for the nonrecursive IRFs.

In the nonrecursive IRF, the response of inflation to a monetary tightening is negative on impact, albeit very small, and, with the exception of the first quarter when it is positive, it stays negative throughout the horizon. Hence, the incidence of a price puzzle is mitigated relative to the recursive IRFs, according to which inflation rises for up to 6 quarters

<sup>8</sup>The kink in the reduced-form representation of the model makes it difficult to approximate the IRFs by local projections on simple nonlinear functions of the data, such as powers or interactions with the regime indicator; see the Supplemental Material for further discussion of this point.

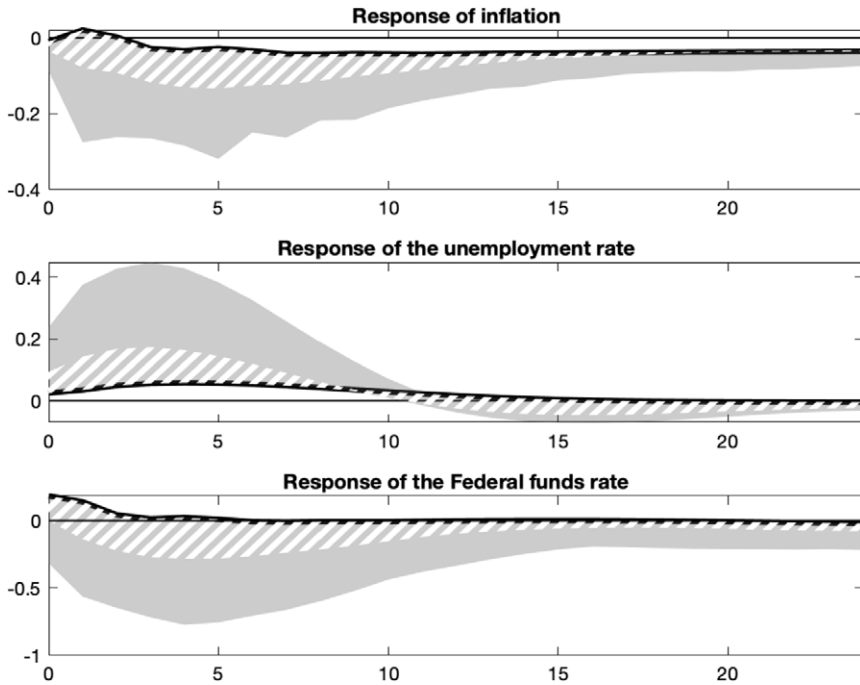


FIGURE 4.—Identified sets of the IRFs to a 25bp monetary policy shock in 2018q3 from a CKSVAR(4) estimated over the period 1960q1 to 2018q2. The shaded area denotes the identified set, the solid line indicates the point-identified IRF under the restriction that unconventional policy is ineffective on impact. The striped area imposes the restriction that the response of the Fed Funds rate on impact should be nonnegative.

after a monetary tightening (9 quarters in the OLS case). Note, however, that the error bands are so wide that they cover (pointwise) most of the recursive IRF, though less so for the OLS one. Turning to the unemployment response, we see that the nonrecursive IRF starts significantly positive on impact (no transmission lag) and peaks much earlier (after 4 quarters) than the recursive IRF (10 quarters). In this case, the recursive IRF is outside the error bands for several quarters (more so for the naive OLS IRF). Finally, the response of the Federal Funds rate to the monetary tightening is less than 1 on impact and generally significantly lower than the recursive IRFs. This is due both to the contemporaneous feedback from inflation and unemployment, as well as the fact that there is a considerable probability of returning to the ZLB, which mitigates the impact of monetary tightening.

Next, I turn to the identified sets of the IRFs that arise when I relax the restriction that unconventional policy is ineffective, that is,  $\lambda$  can be greater than zero. I consider the range of  $\xi = \lambda\zeta \in [0, 1]$ , recalling that  $\lambda$  measures the efficacy of unconventional policy and  $\zeta$  measures the ratio of the reaction function coefficients and shock volatilities in the constrained versus the unconstrained regimes. The shaded areas in Figure 4 report the identified sets without any other restrictions. The striped areas (a subset of the aforementioned identified sets) show the tightening of the identified sets when I impose the additional sign restriction that the contemporaneous effect of the monetary policy shock to the Fed Funds rate should be nonnegative. The bold lines show the IRFs under the (point-identifying) assumption  $\lambda = 0$ . The latter are the same as the nonrecursive point estimates reported in Figure 3.

We observe that the identified set for the IRF of inflation is bounded from above by the limiting case  $\lambda = 0$ . This is also true of the response of the Fed Funds rate. The case  $\lambda = 0$  provides a lower bound on the effect to unemployment only from 0 to 9 quarters. Even though the point estimate of the unemployment response under  $\lambda = 0$  remains positive over all horizons, the identified set includes negative values beyond 10 quarters ahead. We also notice that the identified sets are fairly large, albeit still informative. Interestingly, the identified IRF of the Fed Funds rate includes a range of negative values on impact. These values arise because for values of  $\xi > 0$ , there are generally two solutions for the structural VAR parameters  $\bar{\beta}$ ,  $\bar{\gamma}$  in the equations (35), (39), with one of them inducing such strong responses of inflation and unemployment to the interest rate that the contemporaneous feedback in the policy rule would in fact revert the direct positive effect of the policy shock on the interest rate. If we impose the additional sign restriction that the contemporaneous impact of the policy shock to the Fed Funds rate must be nonnegative, then those values are ruled out and the identified sets become considerably tighter. This is an example of how sign restrictions can lead to tighter partial identification of the IRF.

With an additional assumption on  $\zeta$ , the method can be used to obtain an estimate of the identified set for  $\lambda$ , the measure of the efficacy of unconventional policy. In particular, if we set  $\zeta = 1$ , that is, the reaction function remains the same across the two regimes, then the identified set for  $\lambda$  is  $[0.0.506]$ . In other words, the identified set excludes values of the efficacy of policy beyond 51%, so that, roughly speaking, unconventional policy is at most 51% as effective as conventional one. Note that this estimate does not account for sampling uncertainty and relies crucially on the assumption that the reaction function remains the same across the two regimes. This assumption could be justified by arguing that there is no reason to believe that policy objectives may have shifted over the ZLB period, and that any desired policy stance was feasible over that period. The latter assumption may be questionable. For example, one can imagine that there may be financial and political constraints on the amount of quantitative easing policy makers could do, which may cause them to proceed more cautiously over the ZLB period than over regular times. Within the context of our model, this would be reflected as a flatter policy reaction function over the ZLB period than over the non-ZLB periods, that is, it will correspond to  $\zeta < 1$ . To illustrate the implications of this for the identification of  $\lambda$ , suppose that  $\zeta = 1/2$ , that is, the shadow rate reacts half as fast to shocks during the ZLB period than it does in the non-ZLB period. Then, the identified set for  $\lambda$  would include 1, that is, the data would be consistent with the view that unconventional policy is fully effective. So, under this alternative assumption on  $\zeta$ , the reason we observed a subdued response to policy shocks over the ZLB period is because policy was less active over that period, and policy shocks were smaller, not because unconventional policy was partially ineffective.

As I discussed in the [Introduction](#), it is difficult to make further progress on this issue without further information or additional assumptions. The technical reason is that the scale of the latent regression over the censored sample is not identified, so additional information is required to untangle the structural parameters  $\lambda$  and  $\zeta$  from  $\xi = \lambda\zeta$ . One possibility would be to identify  $\lambda$  from the coefficients on the lags of  $Y_{2t}$  and  $Y_{2t}^*$  by imposing the (overidentifying) restriction that  $Y_{2t-j}$  and  $Y_{2t-j}^*$  appear in the model only via the linear combination  $Y_{2t-j}^{\text{eff}} := \lambda Y_{2t-j}^* + (1 - \lambda) Y_{2t-j}$  for all lags  $j = 1, \dots, p$ , where  $Y_{2t}^{\text{eff}}$  can be interpreted as the effective policy stance. Provided that the coefficients on the lags of  $Y_{2t}^*$  or  $Y_{2t}$  are not all zero, this restriction point-identifies  $\lambda$ , and hence, partially identifies  $\zeta$  from  $\xi$ . One could obtain tighter bounds by using sign restrictions (see [Ikeda et al. \(2020\)](#)), or obtain point identification by using conventional identification schemes. For instance, one can identify  $\bar{\beta}$  directly using external instruments, as in [Gertler and Karadi \(2015\)](#), and hence point-identify  $\xi$  from (39).

5. CONCLUSION

This paper has shown that the ZLB can be used constructively to identify the causal effects of monetary policy on the economy. Identification relies on the (in)efficacy of alternative (unconventional) policies. When unconventional policies are partially effective in mitigating the impact of the ZLB, the causal effects of monetary policy are only partially identified. A general method is proposed to estimate SVARs subject to an occasionally binding constraint. The method can be used to test the efficacy of unconventional policy, modeled via a shadow rate. Application to a core three-equation SVAR with U.S. data suggests that the ZLB is empirically relevant and unconventional policy is only partially effective.

APPENDIX A: PROOFS

A.1. Derivation of Identified Set for Model of Section 2

Using the notation  $\beta^* = \lambda\beta$ , equation (16) can be expressed as

$$\tilde{\beta} = g(\beta)\beta, \quad g(\beta) := \frac{1 - \lambda}{1 - \lambda \frac{\beta(\omega_{12} - \omega_{22}\beta)}{\omega_{11} - \omega_{12}\beta}}. \tag{41}$$

When  $\omega_{12} = 0$ , we have  $g(\beta) = \frac{1-\lambda}{1+\lambda\omega_{22}\beta^2/\omega_{11}} \in (0, 1)$  for all  $\lambda \in (0, 1)$ . Therefore, when  $\tilde{\beta} \neq 0$ , the sign of  $\beta$  is the same as that of  $\tilde{\beta}$  and its magnitude is lower, as stated in (17).

Next, consider  $\omega_{12} \neq 0$ . It is easily seen that  $g(0) = 1 - \lambda$  and  $\lim_{\beta \rightarrow \pm\infty} g(\beta) = 0$ . Moreover,

$$\frac{\partial g}{\partial \beta} = \lambda(1 - \lambda) \frac{\omega_{12}\omega_{22}\beta^2 - 2\omega_{11}\omega_{22}\beta + \omega_{11}\omega_{12}}{(\omega_{11} - \beta\omega_{12} - \beta\lambda\omega_{12} + \beta^2\lambda\omega_{22})^2}.$$

For  $\lambda \in (0, 1)$ , the above derivative function has zeros at  $\omega_{12}\omega_{22}\beta^2 - 2\omega_{11}\omega_{22}\beta + \omega_{11}\omega_{12} = 0$ , which occur at

$$\begin{aligned} \beta_1 &= \frac{\omega_{11}\omega_{22} + \sqrt{\omega_{11}\omega_{22}(\omega_{11}\omega_{22} - \omega_{12}^2)}}{\omega_{12}\omega_{22}}, \\ \beta_2 &= \frac{\omega_{11}\omega_{22} - \sqrt{\omega_{11}\omega_{22}(\omega_{11}\omega_{22} - \omega_{12}^2)}}{\omega_{12}\omega_{22}}, \end{aligned} \quad \text{if } \omega_{12} \neq 0.$$

Now, because  $0 < (\omega_{11}\omega_{22} - \omega_{12}^2) < \omega_{11}\omega_{22}$  implies  $\sqrt{\omega_{11}\omega_{22}(\omega_{11}\omega_{22} - \omega_{12}^2)} < \omega_{11}\omega_{22}$ , we have  $\beta_i < 0, i = 1, 2$ , when  $\omega_{12} < 0$  and  $\beta_i > 0, i = 1, 2$ , when  $\omega_{12} > 0$ .

By symmetry, it suffices to consider only one of the two cases, for example, the case  $\omega_{12} < 0$ . In this case,  $g'(\beta) = \frac{\partial g}{\partial \beta} < 0$  for all  $\beta > 0$  and, since  $g(0) = 1 - \lambda$  and  $g(\infty) = 0$ , it follows that  $g(\beta) \in (0, 1 - \lambda)$  for all  $\beta > 0$ . Thus, from (41), we see that  $\tilde{\beta} < 0$  cannot arise from  $\beta > 0$  when  $\omega_{12} < 0$ . In other words, observing  $\tilde{\beta} < 0$  must mean that  $\beta < 0$ . Moreover, since  $g'(\beta) < 0$  for all  $\beta > \beta_1$  and  $\beta_1 < 0$ , it must be that  $g(\beta) > 0$  for all  $\beta > \beta_1$ , and hence, also for  $\beta_1 < \beta \leq 0$ . At  $\beta < \beta_1$ ,  $g'(\beta) > 0$ , and since  $g'(\beta) < 0$  for all  $\beta < \beta_2 < \beta_1$ , and  $g(-\infty) = 0$ , it has to be that  $g(\beta)$  approaches zero from below as  $\beta \rightarrow -\infty$ , and therefore,  $g(\beta)$  must cross zero at some  $\beta_0 \in (\beta_2, \beta_1)$ , and  $g(\beta) \geq 0$  for

all  $\beta \in [\beta_0, 0]$ . Inspection of (41) shows that  $\beta_0 = \omega_{11}/\omega_{12} = 1/\gamma_0$ , which corresponds to  $\gamma = -\infty$  from (15). Since  $g(\beta) \in [0, 1 - \lambda]$  for all  $\beta \in [\beta_0, 0]$ , and  $\lambda \in (0, 1)$ , it follows from (41) that  $|\tilde{\beta}| \leq |\beta|$ . In other words,  $\tilde{\beta}$  is attenuated relative to the true  $\beta$ .

Finally, we notice that there is a minimum value of  $\tilde{\beta}$  that one can observe under the restriction  $\lambda \in [0, 1]$  (at  $\lambda = 1, \tilde{\beta} = 0$ ). Given the attenuation bias and the fact that  $\tilde{\beta} < 0$  if and only if  $\beta \in [\beta_0, 1]$ , the smallest value of  $\tilde{\beta}$  occurs when  $\lambda = 0$  and  $\beta = \omega_{11}/\omega_{12}$ , so  $\tilde{\beta}_{\min} = \omega_{11}/\omega_{12} = 1/\gamma_0$ . Thus, observing  $\tilde{\beta} < \omega_{11}/\omega_{12}$  and  $\omega_{12} < 0$ , or  $\tilde{\beta}\omega_{12}/\omega_{11} > 1$ , violates the identifying restriction that  $\lambda \geq 0$ , for only with a  $\lambda < 0$  can we get  $g(\beta) > 1$  when  $\beta < 0$  and hence  $\tilde{\beta} < \beta < 0$ .

A.1.1. *Bounds on  $\lambda$*

The bounds on  $\lambda$  are obtained by finding all the values of  $\lambda$  for which equation (41) has a solution for  $\beta$ . This equation implies

$$\beta^2((1 - \lambda)\omega_{12} + \tilde{\beta}\lambda\omega_{22}) - \beta((1 - \lambda)\omega_{11} + \tilde{\beta}(1 + \lambda)\omega_{12}) + \tilde{\beta}\omega_{11} = 0,$$

whose discriminant is the following quadratic function of lambda:

$$D(\lambda) = ((1 - \lambda)\omega_{11} + \tilde{\beta}(1 + \lambda)\omega_{12})^2 - 4\tilde{\beta}\omega_{11}((1 - \lambda)\omega_{12} + \tilde{\beta}\lambda\omega_{22}).$$

Hence, the identified set for  $\lambda$  corresponds to  $S_\lambda = \{\lambda : D(\lambda) \geq 0\}$ . This set is nonempty because  $D(0) \geq 0$ . It can be computed analytically and can take the following three shapes: (i)  $S_\lambda = \mathfrak{R}$  if  $D(\lambda) \geq 0$  for all  $\lambda \in \mathfrak{R}$ ; (ii)  $S_\lambda = (-\infty, \lambda_{lo}] \cup [\lambda_{up}, \infty)$  if  $\omega_{11} - \tilde{\beta}\omega_{12} \neq 0$ , where  $\lambda_{lo} < \lambda_{up}$  are the roots of  $D(\lambda) = 0$ ; and (iii)  $S_\lambda = (-\infty, \lambda_{lo}]$  if  $\omega_{11} - \tilde{\beta}\omega_{12} = 0$ , because  $\omega_{11}^2 - (\frac{\omega_{11}}{\omega_{12}})^2\omega_{12}^2 - 2(\frac{\omega_{11}}{\omega_{12}})\omega_{11}\omega_{12} + 2(\frac{\omega_{11}}{\omega_{12}})^2\omega_{11}\omega_{22} = 2\omega_{11}^2 \frac{\omega_{11}\omega_{22} - \omega_{12}^2}{\omega_{12}^2} > 0$ . If we also impose the restriction  $\lambda \in [0, 1]$ , then the identified set is  $S_\lambda \cap [0, 1]$ .

A.2. *Proof of Proposition 1*

Define  $\bar{A}_{i2} := A_{i2}^* + A_{i2}$ ,  $i = 1, 2$  as the right blocks of  $\bar{A}$  that was defined in (21). Applying Theorem 1 of [Gourieroux, Laffont, and Monfort \(1980\)](#), coherency holds if and only if  $\det \bar{A}$  and  $\det A^*$  have the same sign. Without loss of generality, we can assume that  $A_{11}$  is nonsingular (this can always be achieved by reordering the variables in  $Y_t$ ). From (21), we have  $\det A^* = \det A_{11} \det(A_{22}^* - A_{21}A_{11}^{-1}A_{12}^*)$  and  $\det \bar{A} = \det A_{11} \det(\bar{A}_{22} - A_{21}A_{11}^{-1}\bar{A}_{12})$  ([Lütkepohl \(1996\)](#), p. 50 (6)). The coherency condition can be written as  $\det \bar{A} / \det A^* > 0$ , which, given that  $(\bar{A}_{22} - A_{21}A_{11}^{-1}\bar{A}_{12})$  and  $(A_{22}^* - A_{21}A_{11}^{-1}A_{12}^*)$  are scalars, yields (22).

A.3. *Proof of Proposition 2*

Define  $\bar{A}_{i2} := A_{i2}^* + A_{i2}$ ,  $i = 1, 2$  as the right blocks of  $\bar{A}$  that was defined in (21). Also let  $Y_t^* := (Y_{1t}^*, Y_{2t}^*)'$ . When the coherency condition (22) holds, the solution of (20) exists and is unique. It can be expressed as

$$Y_t^* = \begin{cases} CX_t + C^*X_t^* + u_t & \text{if } D_t = 0, \\ \tilde{C}X_t + \tilde{C}^*X_t^* + \tilde{c}b + \tilde{u}_t & \text{if } D_t = 1, \end{cases} \tag{42}$$

where

$$C = \bar{A}^{-1}B, \quad C^* = \bar{A}^{-1}B^*, \quad u_t = \bar{A}^{-1}\varepsilon_t \tag{43}$$

and

$$\tilde{C} = A^{*-1}B, \quad \tilde{C}^* = A^{*-1}B^*, \quad \tilde{c} = -A^{*-1} \begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix} b, \quad \tilde{u}_t = A^{*-1}\varepsilon_t. \tag{44}$$

Using the partitioned inverse formula, we obtain

$$\begin{aligned} \tilde{C}_1 &= (A_{11} - A_{12}^*A_{22}^{*-1}A_{21})^{-1}(B_1 - A_{12}^*A_{22}^{*-1}B_2), \\ \tilde{C}_2 &= (A_{22}^* - A_{21}A_{11}^{-1}A_{12}^*)^{-1}(B_2 - A_{21}A_{11}^{-1}B_1), \end{aligned}$$

and

$$\begin{aligned} C_1 &= (A_{11} - \bar{A}_{12}\bar{A}_{22}^{-1}A_{21})^{-1}(B_1 - \bar{A}_{12}\bar{A}_{22}^{-1}B_2), \\ C_2 &= (\bar{A}_{22} - A_{21}A_{11}^{-1}\bar{A}_{12})^{-1}(B_2 - A_{21}A_{11}^{-1}B_1). \end{aligned}$$

Solving the latter for  $B_1$  and  $B_2$  yields

$$B_1 = A_{11}C_1 + \bar{A}_{12}C_2, \quad \text{and} \quad B_2 = \bar{A}_{22}C_2 + A_{21}C_1.$$

Thus,

$$\begin{aligned} \tilde{C}_1 &= C_1 + (A_{11} - A_{12}^*A_{22}^{*-1}A_{21})^{-1}(\bar{A}_{12} - A_{12}^*A_{22}^{*-1}\bar{A}_{22})C_2 = C_1 - \tilde{\beta}C_2 \quad \text{and} \\ \tilde{C}_2 &= (A_{22}^* - A_{21}A_{11}^{-1}A_{12}^*)^{-1}(A_{22}^* - A_{21}A_{11}^{-1}A_{12}^* + A_{22} - A_{21}A_{11}^{-1}A_{12})C_2 = \kappa C_2, \end{aligned}$$

where  $\kappa$  is given in (22). The exact same derivations apply to  $\tilde{C}^*$ , that is,

$$\tilde{C}_1^* = C_1^* - \tilde{\beta}C_2^* \quad \text{and} \quad \tilde{C}_2^* = \kappa C_2^*.$$

Next,

$$\begin{aligned} \tilde{c}_1 &= (A_{11} - A_{12}^*A_{22}^{*-1}A_{21})^{-1}(A_{12}^*A_{22}^{*-1}A_{22} - A_{12})b = \tilde{\beta}b \quad \text{and} \\ \tilde{c}_2 &= -\frac{A_{22} - A_{21}A_{11}^{-1}A_{12}}{A_{22}^* - A_{21}A_{11}^{-1}A_{12}^*}b = (1 - \kappa)b. \end{aligned}$$

Finally,  $\tilde{u}_t = A^{*-1}\bar{A}u_t = (\tilde{u}'_{1t}, \tilde{u}'_{2t})'$ , where

$$\tilde{u}_{1t} = u_{1t} - \tilde{\beta}u_{2t} \quad \text{and} \quad \tilde{u}_{2t} = \kappa u_{2t}.$$

Substituting back into (42), the reduced-form model for  $Y_{1t}$  becomes

$$\begin{aligned} Y_{1t} &= (1 - D_t)(C_1X_t + C_1^*X_t^* + u_{1t}) \\ &\quad + D_t((C_1 - \tilde{\beta}C_2)X_t + (C_1^* - \tilde{\beta}C_2^*)X_t^* + u_{1t} - \tilde{\beta}u_{2t}), \end{aligned} \tag{45}$$

and for  $Y_{2t}^*$  it is

$$Y_{2t}^* = C_2 X_t + C_2^* X_t^* + u_{2t} - (1 - \kappa) D_t (C_2 X_t + C_2^* X_t^* + u_{2t} - b). \tag{46}$$

Next, define

$$\tilde{Y}_{2t}^* := C_2 X_t + C_2^* X_t^* + u_{2t}, \tag{47}$$

and rewrite (46) as

$$\begin{aligned} Y_{2t}^* &= \tilde{Y}_{2t}^* - (1 - \kappa) D_t (\tilde{Y}_{2t}^* - b) \\ &= (1 - D_t) \tilde{Y}_{2t}^* + D_t (\kappa \tilde{Y}_{2t}^* + (1 - \kappa) b). \end{aligned} \tag{48}$$

Let  $q = \dim X_t$  denote the number of elements of  $X_t$  and define, for each  $i = 1, 2$ ,

$$\bar{C}_{ij} = \begin{cases} C_{ij}, & j \in \{1, q\} : X_{ij} \neq Y_{2,t-s} \text{ for all } s \in \{1, p\}, \\ C_{ij} + C_{is}^*, & j \in \{1, q\} : X_{ij} = Y_{2,t-s}, \text{ for some } s \in \{1, p\}. \end{cases} \tag{49}$$

In other words,  $\bar{C}$  contains the original coefficients on all the regressors other than the lags of  $Y_{2t}$ , while the coefficients on the lags of  $Y_{2t}$  are augmented by the corresponding coefficients of the lags of  $Y_{2t}^*$ . For example, if  $p = 1$  and there are no other exogenous regressors  $X_{0t}$ , then, for  $i = 1, 2$ ,

$$C_i X_t + C_i^* X_t^* = C_{i1} Y_{1t-1} + C_{i2} Y_{2t-1} + C_i^* Y_{2t-1}^*,$$

so  $\bar{C}_i = (C_{i1}, C_{i2} + C_i^*)$ . Using (49), we can rewrite (47) as

$$\tilde{Y}_{2t}^* = \bar{C}_2 X_t + C_2^* \min(X_t^* - b, 0) + u_{2t}. \tag{50}$$

Now, observe that

$$\min(Y_{2t}^* - b, 0) = D_t (Y_{2t}^* - b) = \kappa D_t (\tilde{Y}_{2t}^* - b) = \kappa \min(\tilde{Y}_{2t}^* - b, 0).$$

So, letting  $\tilde{X}_t^*$  denote the lags of  $\tilde{Y}_{2t}^*$ , we have  $\min(X_t^* - b, 0) = \kappa \min(\tilde{X}_t^* - b, 0)$ , and consequently,

$$C^* \min(X_t^* - b, 0) = \bar{C}^* \min(\tilde{X}_t^* - b, 0),$$

where  $\bar{C}^* = \kappa C^*$ . Now, from (50) we have

$$\tilde{Y}_{2t}^* = \bar{C}_2 X_t + \bar{C}_2^* \min(\tilde{X}_t^* - b, 0) + u_{2t}.$$

Recall the definition of  $\bar{Y}_{2t}^*$  in (26):

$$\bar{Y}_{2t}^* := \bar{C}_2 X_t + \bar{C}_2^* \bar{X}_t^* + u_{2t},$$

where  $\bar{X}_t^* := (\bar{x}_{t-1}, \dots, \bar{x}_{t-p})'$ , and  $\bar{x}_t := \min(\bar{Y}_{2t}^* - b, 0)$ , with initial conditions  $\bar{x}_{-s} = \kappa^{-1} \min(Y_{2,-s}^* - b, 0)$ ,  $s = 0, \dots, p - 1$ . It follows that  $\min(\tilde{X}_t^* - b, 0) = \bar{X}_t^*$  for all  $t \geq 1$ , so



that  $\tilde{Y}_{2t}^* = \bar{Y}_{2t}^*$ . Substituting  $\bar{Y}_2^*$  for  $\tilde{Y}_2^*$  in (48), we get (27). Using the reparameterization (49) and the relationship between  $X_t^*$  and  $\bar{X}_t^*$  in (45), we obtain (24).

Finally, from equation (46), it follows that the event  $Y_{2t}^* < b$  is equivalent to

$$b + \kappa(C_2 X_t + C_2^* X_t^* + u_{2t} - b) < b,$$

which, since  $\kappa > 0$  by the coherency condition (22), is equivalent to

$$u_{2t} < b - C_2 X_t - C_2^* X_t^*. \tag{51}$$

Using the definition (26), and (49), the inequality (51) can be written as  $\bar{Y}_{2t}^* < b$ , which establishes (25).

*Comment:* Note that  $\kappa$  appears in the reduced form only multiplicatively with  $C^*$ , so  $\kappa$  and  $C^*$  are not separately identified; only  $\bar{C}^* = \kappa C^*$  is. The reparameterization from  $C$  to  $\bar{C}$  is convenient because  $\bar{C}$  is identified independently of  $\kappa$ , while  $C$ ,  $C^*$ , and  $\kappa$  are not separately identified.

#### A.4. Proof of Proposition 3

We solve  $u_t = \bar{A}^{-1} \varepsilon_t$  using the partitioned inverse formula to get

$$u_{1t} = (A_{11} - \bar{A}_{12} \bar{A}_{22}^{-1} A_{21})^{-1} (\varepsilon_{1t} - \bar{A}_{12} \bar{A}_{22}^{-1} \varepsilon_{2t}), \tag{52}$$

$$u_{2t} = (\bar{A}_{22} - A_{21} A_{11}^{-1} \bar{A}_{12})^{-1} (\varepsilon_{2t} - A_{21} A_{11}^{-1} \varepsilon_{1t}). \tag{53}$$

Using the definitions

$$\begin{aligned} \bar{\beta} &:= -A_{11}^{-1} \bar{A}_{12}, & \bar{\gamma} &:= -\bar{A}_{22}^{-1} A_{21}, \\ \bar{\varepsilon}_{1t} &:= A_{11}^{-1} \varepsilon_{1t}, & \bar{\varepsilon}_{2t} &:= \bar{A}_{22}^{-1} \varepsilon_{2t}, \end{aligned}$$

we can rewrite (52)–(53) as (32)–(33).

Note that

$$\begin{aligned} \bar{\varepsilon}_{1t} &= A_{11}^{-1} (A_{11} u_{1t} + \bar{A}_{12} u_{2t}) = u_{1t} - \bar{\beta} u_{2t}, \\ \bar{\varepsilon}_{2t} &= \bar{A}_{22}^{-1} (A_{21} u_{1t} + \bar{A}_{22} u_{2t}) = -\bar{\gamma} u_{1t} + u_{2t}, \end{aligned}$$

so,

$$\begin{aligned} \text{var}(\bar{\varepsilon}_{1t}) &= (I_{k-1}, -\bar{\beta}) \Omega (I_{k-1}, -\bar{\beta})', \\ \text{var}(\bar{\varepsilon}_{2t}) &= (-\bar{\gamma}, 1) \Omega (-\bar{\gamma}, 1)', \end{aligned}$$

and

$$\begin{aligned} \text{cov}(\bar{\varepsilon}_{1t}, \bar{\varepsilon}_{2t}) &= (I_{k-1}, -\bar{\beta}) \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega'_{12} & \Omega_{22} \end{pmatrix} (-\bar{\gamma}, 1)' \\ &= -(\Omega_{11} - \bar{\beta} \Omega'_{12}) \bar{\gamma}' + \Omega_{12} - \bar{\beta} \Omega_{22} = 0. \end{aligned}$$

TABLE B.I  
PARAMETER NOTATION IN REPORTED SIMULATION RESULTS.

Mnemonic	Description
$\tau$	st. dev. of reduced-form error $u_{2t}$ in $Y_{2t}$ (constrained variable)
Eq. 3	reduced-form equation for $Y_{2t}$
Eq. j	red.-form equation for $Y_{1j,t}$ , $j = 1, 2$ (unconstrained variables)
$\tilde{\beta}_j$	coefficient on kink in eq. j
eq. i Y1j_1	coefficient of $Y_{1j,t-1}$ in eq. i
eq. i Y2_1	coefficient of $Y_{2,t-1}$ in eq. i
eq. i lY2_1	coefficient of $\min(Y_{2,t-1}^* - b, 0)$ in eq. i
$\delta_j$	coefficient of regression of $u_{1j,t}$ (red.-form error in eq. j) on $u_{2t}$
Ch_ij	(i,j) element of Choleski factor of $\Omega_{1,2}$

The last equation identifies  $\bar{\gamma}$  given  $\bar{\beta}$ . Specifically,

$$\bar{\gamma}' = (\Omega_{11} - \bar{\beta}\Omega'_{12})^{-1}(\Omega_{12} - \bar{\beta}\Omega_{22}).$$

APPENDIX B: NUMERICAL RESULTS

This section provides Monte Carlo evidence on the finite-sample properties of the proposed estimators and tests. The data generating process (DGP) is a trivariate VAR(1), given by equations (18) and (19). I consider three different DGPs corresponding to the CKSVAR, KSVAR, and CSVAR models, respectively. In all three DGPs, the following parameters are set to the same values: the contemporaneous coefficients are  $A_{11} = I_2$ ,  $A_{12} = A_{12}^* = 0_{2 \times 1}$ ,  $A_{22}^* = 1$ , and  $A_{22} = 0$ ; the intercepts are set to zero,  $B_{10} = 0_{2 \times 1}$  and  $B_{20} = 0$ ; the coefficients on the lags are  $B_{1,1} = (\rho I_2, 0)$ ,  $B_{1,1}^* = 0_{2 \times 1}$ ,  $B_{2,1} = (0_{1 \times 2}, B_{22,1})$ , with  $\rho = 0.5$ . Finally, each of the three DGPs is determined as follows. DGP1:  $B_{22,1} = B_{2,1}^* = 0$  (both KSVAR and CSVAR restrictions hold, since lags of  $Y_{2,t}$  and  $Y_{2,t}^*$  all have zero coefficients); DGP2:  $B_{22,1} = \rho$ ,  $B_{2,1}^* = 0$  (KSVAR restrictions hold but CSVAR restrictions do not); DGP3:  $B_{22,1} = 0$ ,  $B_{2,1}^* = \rho$  (CSVAR restrictions hold but KSVAR restrictions do not). The setting of the autoregressive coefficient  $\rho = 0.5$  leads to a lower degree of persistence than is typically observed in macro data (e.g., in the Stock and Watson (2001) application, the three largest roots are 0.97, 0.97, and 0.8), because I want to avoid confounding any possible finite-sample issues arising from the ZLB with well-known problems of bias and size distortion due to strong persistence (near unit roots) in the data. Finally, the bound on  $Y_{2t}$  is set to  $b = 0$ , the sample size is  $T = 250$ , the initial conditions are set to 0, and the number of Monte Carlo replications is 1000. In all cases, the CKSVAR and CSVAR likelihoods are computed using SIS with  $R = 1000$  particles. The notation for the reported parameters is given in Table B.I.

Figure B.1 reports the sampling distribution of the ML estimators of the reduced-form parameters in Proposition 2 for the CKSVAR model under DGP1. The results for the KSVAR and CSVAR models, which are also correctly specified under DGP1, are omitted because they are entirely analogous. The sampling densities appear to be very close to the superimposed Normal approximations, indicating that the Normal asymptotic approximation is fairly accurate.

Table B.II reports moments of the sampling distributions of the above mentioned estimators for the CKSVAR model. Again, the results for the KSVAR and CSVAR models are entirely analogous and are therefore omitted. We notice no discernible biases. Results

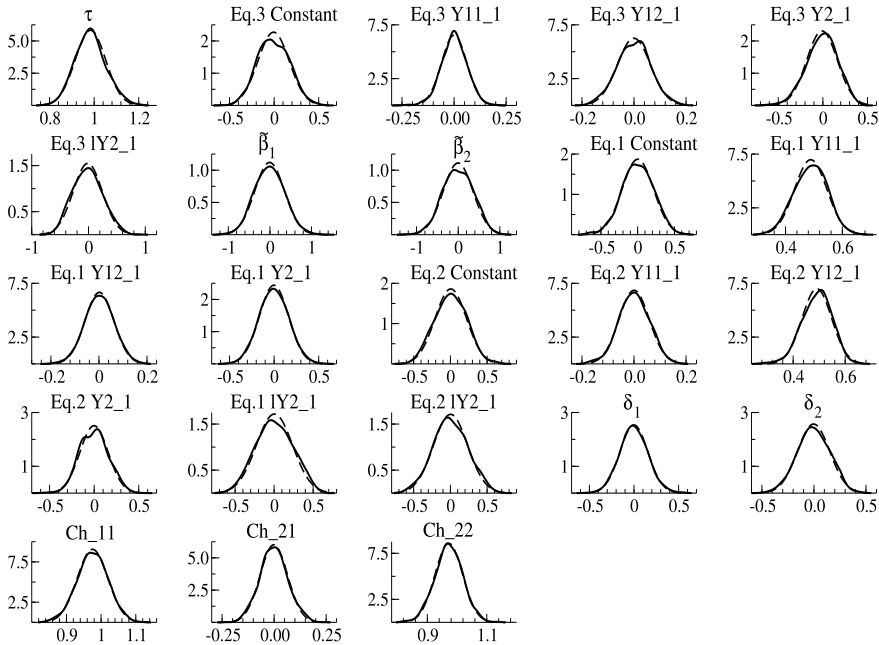


FIGURE B.1.—Sampling densities of ML estimators of reduced-form coefficients of CKSVAR(1) under DGP1 (solid lines) and approximating Normal densities (dashed lines).  $T = 250$ , 1000 Monte Carlo replications. Parameter names described in Table B.I.

with  $T = 100$  and  $T = 1000$  given in the Supplemental Material indicate that the RMSE declines at rate  $\sqrt{T}$  in accordance with asymptotic theory. It is noteworthy that the estimators of  $\tilde{\beta}$  have substantially larger RMSE than the estimators of the other parameters.

The DGPs in the previous simulations have the property that the frequency of the ZLB regime is around 50%. The Supplemental Material reports simulation results with a slight modification to the DGPs to match the frequency in the sample of the empirical application in the paper (11%). The results are very similar to those reported in Figure B.1 and Table B.II: the Normal approximation of the sampling distribution of the ML estimator remains very accurate and the bias is negligible. The only difference is that the standard deviation of  $\tilde{\beta}$  is larger.

Next, I turn to the properties of the LR test of KSVAR against CKSVAR and CSVAR against CKSVAR. The former hypothesis involves three restrictions (exclusion of the latent lag  $Y_{2,t-1}^*$  from each of the three equations), so the LR statistic is asymptotically distributed as  $\chi_3^2$  under the null. The latter hypothesis involves five restrictions (exclusion of the observed lag  $Y_{2,t-1}$  from each of the three equations, plus  $\tilde{\beta} = 0$ ), and the LR statistic is asymptotically distributed as  $\chi_5^2$ . Table B.III reports the rejection frequencies of the LR tests for each of the two hypotheses in each of the three DGPs at three significance levels: 10%, 5%, and 1%. In addition to the asymptotic tests, I also report the rejection frequency of the tests using parametric bootstrap critical values. The parametric bootstrap uses draws of Normal errors and the estimated reduced-form parameters to generate the bootstrap samples of  $Y_t$  and  $Y_{2t}^*$ . The Monte Carlo rejection frequencies are computed using the “warp-speed” method of Giacomini, Politis, and White (2013). Note that both null hypotheses hold under DGP1, but only the KSVAR is valid under DGP2 and only

TABLE B.II  
MOMENTS OF SAMPLING DISTRIBUTION OF ML ESTIMATORS OF THE PARAMETERS OF CKSVAR(1).<sup>a</sup>

ML-CKSVAR	True	Mean	Bias	sd	RMSE
$\tau$	1.000	0.983	-0.017	0.068	0.070
Eq.3 Constant	0.000	-0.006	-0.006	0.175	0.176
Eq.3 Y11_1	0.000	0.000	0.000	0.061	0.061
Eq.3 Y12_1	0.000	-0.000	-0.000	0.064	0.064
Eq.3 Y2_1	0.000	-0.010	-0.010	0.173	0.173
Eq.3 1Y2_1	0.000	-0.013	-0.013	0.258	0.258
$\tilde{\beta}_1$	-0.000	-0.008	-0.008	0.356	0.356
$\tilde{\beta}_2$	0.000	-0.004	-0.004	0.359	0.359
Eq.1 Constant	0.000	0.002	0.002	0.213	0.213
Eq.1 Y11_1	0.500	0.489	-0.011	0.057	0.058
Eq.1 Y12_1	0.000	0.002	0.002	0.060	0.060
Eq.1 Y2_1	0.000	-0.002	-0.002	0.163	0.163
Eq.2 Constant	0.000	0.005	0.005	0.214	0.214
Eq.2 Y11_1	0.000	0.000	0.000	0.058	0.058
Eq.2 Y12_1	0.500	0.491	-0.009	0.057	0.058
Eq.2 Y2_1	0.000	-0.001	-0.001	0.159	0.159
Eq.1 1Y2_1	0.000	0.005	0.005	0.232	0.232
Eq.2 1Y2_1	0.000	0.002	0.002	0.233	0.233
$\delta_1$	0.000	-0.001	-0.001	0.157	0.157
$\delta_2$	0.000	-0.004	-0.004	0.155	0.155
Ch_11	1.000	0.975	-0.025	0.044	0.051
Ch_21	0.000	-0.001	-0.001	0.066	0.066
Ch_22	1.000	0.972	-0.028	0.046	0.054

<sup>a</sup>Computed under DGP1 with  $T = 250$  using 1000 MC replications. Parameter names described in Table B.I.

the CSVAR is valid under DGP3. For convenience, I indicate the rejection frequencies under the alternative in bold in the table.

There is evidence that the LR tests reject too often under  $H_0$  relative to their nominal level when we use asymptotic critical values. Moreover, the size distortions are very similar across null hypotheses and DGPs. Unreported results show that size distortion eventually disappears as the sample gets large, but this level of over-rejection is unsatisfactory at  $T = 250$ , which is a typical sample size in macroeconomic applications. The paramet-

TABLE B.III  
REJECTION FREQUENCIES OF LR TESTS OF  $H_0$  AGAINST  $H_1$  ACROSS DIFFERENT DGPs.<sup>a</sup>

Sign. Level	$H_0$ : KSVAR, $H_1$ : CKSVAR			$H_0$ : CSVAR, $H_1$ : CKSVAR			
	10%	5%	1%	10%	5%	1%	
DGP1	asymptotic	0.173	0.093	0.020	0.155	0.084	0.022
	bootstrap	0.107	0.045	0.005	0.109	0.052	0.012
DGP2	asymptotic	0.149	0.080	0.018	<b>0.319</b>	<b>0.206</b>	<b>0.073</b>
	bootstrap	0.117	0.050	0.011	<b>0.242</b>	<b>0.141</b>	<b>0.041</b>
DGP3	asymptotic	<b>0.587</b>	<b>0.454</b>	<b>0.244</b>	0.141	0.067	0.019
	bootstrap	<b>0.471</b>	<b>0.365</b>	<b>0.194</b>	0.103	0.053	0.018

<sup>a</sup>Computed using 1000 Monte Carlo replications,  $T = 250$ . The asymptotic tests use  $\chi^2_3$  and  $\chi^2_5$  critical values for KSVAR and CSVAR resp. The bootstrap rej. frequencies were computed using the warp-speed method of Giacomini, Politis, and White (2013). Bold numbers indicate that the rejection frequencies were computed under  $H_1$  (power).

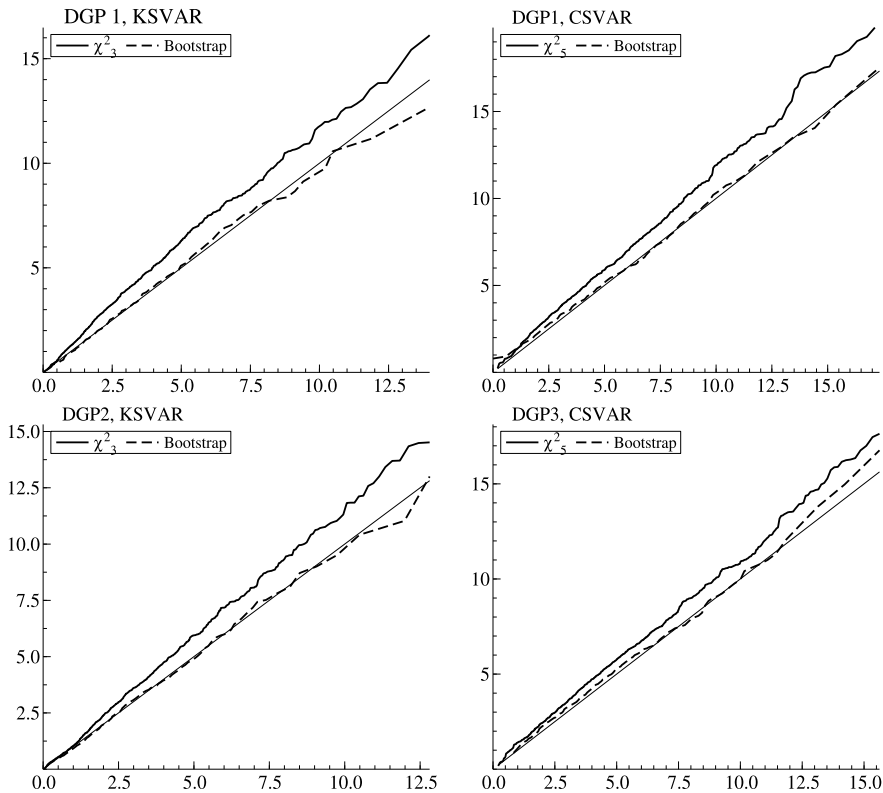


FIGURE B.2.—QQ plots of the sampling distribution under the null hypothesis of LR statistics of KSVAR against CKSVAR (left) and CSVAR against CKSVAR (right). Solid (dashed) lines plot quantiles against asymptotic  $\chi^2$  (bootstrap) approximation. Computed for  $T = 250$  using 1000 Monte Carlo replications.

ric bootstrap appears to do a remarkably good job at correcting the size of the tests. In all cases considered, the parametric bootstrap rejection frequency is not significantly different from the nominal level when the null hypothesis holds (all but the numbers in bold in the Table). To shed further light on this issue, Figure B.2 reports the QQ plots of the sampling distributions of the two LR statistics against their asymptotic and parametric bootstrap approximations for all three DGPs under the null hypothesis. The sampling distributions of the LR statistics stochastically dominate their asymptotic approximations, but the bootstrap approximations are quite accurate.

Finally, the rejection frequencies highlighted in bold in Table B.III correspond to the power of the tests against two very similar deviations from the null hypothesis. The numbers on the left under DGP3 show the power of the test to reject the KSVAR specification under the alternative at which the coefficient on the latent lag  $B_{2,1}^* = 0.5$ . Similarly, the bold numbers on the right give the power of rejecting CSVAR against the alternative where the coefficient on the observed lag  $B_{22,1} = 0.5$ . Since the lower bound is set to zero, and the sample contains about 50% of observations at the ZLB, the two deviations from the null are of equal magnitude. Yet, we notice the LR test is significantly more powerful against the KSVAR than against the CSVAR. This could be because CSVAR imposes more restrictions than the KSVAR, so one would expect it to have lower power than the KSVAR against similar deviations from the null.

## REFERENCES

- AMEMIYA, T. (1974): "Multivariate Regression and Simultaneous Equation Models When the Dependent Variables Are Truncated Normal," *Econometrica*, 42 (6), 999–1012. [2856,2859]
- ARUOBA, S. B., P. CUBA-BORDA, K. HIGA-FLORES, F. SCHORFHEIDE, AND S. VILLALVAZO (2020): "Piecewise-Linear Approximations and Filtering for DSGE Models With Occasionally Binding Constraints," Technical report. [2856]
- ARUOBA, S. B., P. CUBA-BORDA, AND F. SCHORFHEIDE (2017): "Macroeconomic Dynamics Near the ZLB: A Tale of Two Countries," *The Review of Economic Studies*, 85 (1), 87–118. [2856]
- ARUOBA, S. B., F. SCHORFHEIDE, AND S. VILLALVAZO (2020): "SVARs With Occasionally-Binding Constraints," Technical report. [2857]
- BLUNDELL, R., AND R. J. SMITH (1994): "Coherency and Estimation in Simultaneous Models With Censored or Qualitative Dependent Variables," *Journal of Econometrics*, 64 (1-2), 355–373. [2856,2859,2868]
- BLUNDELL, R. W., AND R. J. SMITH (1989): "Estimation in a Class of Simultaneous Equation Limited Dependent Variable Models," *The Review of Economic Studies*, 56 (1), 37–57. [2859]
- CHEN, H., V. CÚRDIA, AND A. FERRERO (2012): "The Macroeconomic Effects of Large-Scale Asset Purchase Programmes," *The Economic Journal*, 122, F289–F315. [2858]
- DEBORTOLI, D., J. GALI, AND L. GAMBETTI (2019): "On the Empirical (Ir) Relevance of the Zero Lower Bound Constraint," in *NBER Macroeconomics Annual 2019*, Vol. 34. University of Chicago Press. [2856, 2858,2859,2864]
- EGGERTSSON, G. B., AND M. WOODFORD (2003): "Zero Bound on Interest Rates and Optimal Monetary Policy," *Brookings Papers on Economic Activity*, 2003 (1), 139–233. [2855]
- FERNÁNDEZ-VILLAVARDE, J., G. GORDON, P. GUERRÓN-QUINTANA, AND J. F. RUBIO-RAMÍREZ (2015): "Nonlinear Adventures at the Zero Lower Bound," *Journal of Economic Dynamics and Control*, 57, 182–204. [2856]
- GERTLER, M., AND P. KARADI (2015): "Monetary Policy Surprises, Credit Costs, and Economic Activity," *American Economic Journal: Macroeconomics*, 7 (1), 44–76. [2874]
- GIACOMINI, R., D. N. POLITIS, AND H. WHITE (2013): "A Warp-Speed Method for Conducting Monte Carlo Experiments Involving Bootstrap Estimators," *Econometric Theory*, 29 (3), 567–589. [2881,2882]
- GOURIEROUX, C., J. LAFFONT, AND A. MONFORT (1980): "Coherency Conditions in Simultaneous Linear Equation Models With Endogenous Switching Regimes," *Econometrica*, 48 (3), 675–695. [2860,2876]
- GREENE, W. H. (1993): *Econometric Analysis*. New York: MacMillan. [2856]
- GUERRIERI, L., AND M. IACOVIELLO (2015): "OccBin: A Toolkit for Solving Dynamic Models With Occasionally Binding Constraints Easily," *Journal of Monetary Economics*, 70, 22–38. [2856]
- HAYASHI, F., AND J. KOEDA (2019): "Exiting From Quantitative Easing," *Quantitative Economics*, 10, 1069–1107. [2856,2857]
- HECKMAN, J. J. (1978): "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 46 (4), 931–959. [2861]
- (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47 (1), 153–161. [2861]
- IKEDA, D., S. LI, S. MAVROEIDIS, AND F. ZANETTI (2020): "Testing the Effectiveness of Unconventional Monetary Policy in Japan and the United States," Discussion paper 2020-E-10, Institute for Monetary and Economic Studies, Bank of Japan, Available at <https://www.imes.boj.or.jp/research/papers/english/20-E-10.pdf>. [2858,2874]
- KOOP, G., M. H. PESARAN, AND S. M. POTTER (1996): "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics*, 74 (1), 119–147. [2866,2872]
- KULISH, M., J. MORLEY, AND T. ROBINSON (2017): "Estimating DSGE Models With Zero Interest Rate Policy," *Journal of Monetary Economics*, 88 (C), 35–49. [2856]
- LEE, L.-F. (1976): "Multivariate Regression and Simultaneous Equations Models With Some Dependent Variables Truncated," Discussion paper 76-79, University of Minnesota, Minneapolis, USA. [2856,2859]
- (1999): "Estimation of Dynamic and ARCH Tobit Models," *Journal of Econometrics*, 92 (2), 355–390. [2856,2865,2866,2869]
- LEWBEL, A. (2007): "Coherency and Completeness of Structural Models Containing a Dummy Endogenous Variable," *International Economic Review*, 48 (4), 1379–1392. [2860]
- LIEBSCHER, E. (2005): "Towards a Unified Approach for Proving Geometric Ergodicity and Mixing Properties of Nonlinear Autoregressive Processes," *Journal of Time Series Analysis*, 26 (5), 669–689. [2869]
- LIU, P., K. THEODORIDIS, H. MUMTAZ, AND F. ZANETTI (2019): "Changing Macroeconomic Dynamics at the Zero Lower Bound," *Journal of Business & Economic Statistics*, 37 (3), 391–404. [2856]
- LÜTKEPOHL, H. (1996): *Handbook of Matrices*. England: Wiley. [2876]
- MAGNUSSON, L. M., AND S. MAVROEIDIS (2014): "Identification Using Stability Restrictions," *Econometrica*, 82 (5), 1799–1851. [2856]

- MALIK, S., AND M. K. PITT (2011): "Particle Filters for Continuous Likelihood Evaluation and Maximisation," *Journal of Econometrics*, 165 (2), 190–209. [2869]
- MAVROEIDIS, S. (2021): "Supplement to 'Identification at the Zero Lower Bound'," *Econometrica Supplemental Material*, 89, <https://doi.org/10.3982/ECTA17388>. [2857]
- NELSON, F., AND L. OLSON (1978): "Specification and Estimation of a Simultaneous-Equation Model With Limited Dependent Variables," *International Economic Review*, 19 (3), 695–709. [2859,2860]
- NEWBY, W. K., AND D. MCFADDEN (1994): "Large Sample Estimation and Hypothesis Testing," in *The Handbook of Econometrics*, Vol. 4, ed. by R. F. Engle and D. McFadden. North-Holland, 2111–2245. [2869]
- PITT, M. K., AND N. SHEPHARD (1999): "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, 94 (446), 590–599. [2856]
- REIFSCHEIDER, D., AND J. C. WILLIAMS (2000): "Three Lessons for Monetary Policy in a Low-Inflation Era," *Journal of Money, Credit and Banking*, 32 (4), 936–966. [2858,2864]
- RIGOBON, R. (2003): "Identification Through Heteroskedasticity," *The Review of Economics and Statistics*, 85 (4), 777–792. [2856]
- ROSSI, B. (2019): "Identifying and Estimating the Effects of Unconventional Monetary Policy: How to Do It and What Have We Learned?" Discussion Paper DP14064, CEPR. [2857]
- SMITH, R. J., AND R. W. BLUNDELL (1986): "An Exogeneity Test for a Simultaneous Equation Tobit Model With an Application to Labor Supply," *Econometrica*, 54 (3), 679–685. [2859]
- STOCK, J. H., AND M. W. WATSON (2001): "Vector Autoregressions," *Journal of Economic Perspectives*, 15 (4), 101–115. [2857,2869,2870,2872]
- SWANSON, E. T., AND J. C. WILLIAMS (2014): "Measuring the Effect of the Zero Lower Bound on Medium- and Longer-Term Interest Rates," *American Economic Review*, 104 (10), 3154–3185. [2859]
- TOBIN, J. (1958): "Estimation of Relationships for Limited Dependent Variables," *Econometrica*, 26 (1), 24–36. [2856,2859]
- WU, J. C., AND F. D. XIA (2016): "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound," *Journal of Money, Credit and Banking*, 48 (2-3), 253–291. [2857]
- WU, J. C., AND J. ZHANG (2019): "A Shadow Rate New Keynesian Model," *Journal of Economic Dynamics and Control*, 107, 103728. [2859]

---

Co-editor Ulrich K. Müller handled this manuscript.

Manuscript received 10 June, 2019; final version accepted 5 May, 2021; available online 14 May, 2021.