SVARs With Occasionally-Binding Constraints

S. Borağan Aruoba  Frank Schorfheide  Sergio Villalvazo
University of Maryland  University of Pennsylvania  University of Pennsylvania
CEPR, NBER, PIER

Current Version: May 31, 2020

Abstract

We develop a structural VAR in which an occasionally-binding constraint generates censoring of one of the dependent variables. Once the censoring mechanism is triggered, we allow some of the coefficients for the remaining variables to change. By imposing that the regression functions are continuous at the censoring point, we can show that under some mild parameter restrictions delivers a unique reduced form. In our application the occasionally-binding constraint is the effective lower bound on nominal interest rates. According to our estimates based on U.S. data, once the ELB becomes binding, in addition to the censoring of the nominal interest rate, the coefficients in the inflation equation change. This coefficient switch translates into a change of the inflation responses to (unconventional) monetary policy shocks and demand shocks. Our results suggests that the presence of the ELB is indeed empirically relevant for the propagation of shocks.

JEL CLASSIFICATION: C11, C22, C34, E32, E52

KEY WORDS: Bayesian inference, effective lower bound, limited dependent variables, sequential Monte Carlo methods, structural VARs.

* Correspondence: B. Aruoba (aruoba@umd.edu), Department of Economics, University of Maryland, College Park, MD 20742. F. Schorfheide (schorf@ssc.upenn.edu), S. Villalvazo (vsergio@sas.upenn.edu): Department of Economics, University of Pennsylvania, 133 S. 36th Street, Philadelphia, PA 19104. Aruoba and Schorfheide gratefully acknowledge financial support from the National Science Foundation under Grant SES 1851634.
1 Introduction

Dynamic stochastic general equilibrium (DSGE) models are widely used in central banks, by regulators, and in academia to study the effects of monetary and macroprudential policies and the propagation of shocks in the macro economy. The most recent vintage of these models involves occasionally-binding constraints arising from, for instance, an effective lower bound (ELB) on nominal interest rates. Agents’ decision rules in these models typically exhibit “kinks,” meaning the elasticity of the choice variables, say output or prices, with respect to the underlying state variables changes drastically when the constraint becomes binding. We use this observation in Aruoba, Cuba-Borda, Higa-Flores, Schorfheide, and Villalvazo (2020), henceforth ACHSV, to approximate such decision rules through piecewise-linear and continuous (PLC) functions. In turn, the model solution has the form of a vector autoregression (VAR) with a censored dependent variable and regime switching coefficients. The regime shift is endogenous and it is linked to the censoring mechanism. For instance, in a model with an ELB constraint, the VAR coefficients switch, once the interest rate becomes zero.

The goal of this paper is to develop a structural VAR (SVAR) that mimics the PLC-DSGE model solution, but can be used independently of an optimization-based structural model to study the propagation of shocks in settings where one (or more) observable is subject to an occasionally-binding constraint. Throughout this paper, we focus on an application that features nominal interest rates that are constrained by an ELB.\textsuperscript{1} An important empirical question is whether the propagation of shocks works differently when the economy reaches the ELB.

For instance, several authors have argued based on DSGE models that unanticipated changes in government spending generate larger multipliers when the nominal interest rate is constrained; see Eggertsson (2011) and Christiano, Eichenbaum, and Rebelo (2011). In a more densely parameterized model such as SVARs, these effects are empirically difficult to measure because for many countries we only have a few years of ELB observations available. This makes it difficult to obtain precise estimates. Our proposed model is able to avoid this

\textsuperscript{1}We use the concept of effective instead of zero lower bound (ZLB). In a DSGE model that explicitly models money demand, such as the one in Aruoba and Schorfheide (2011), an interest rate less than zero means monetary equilibrium ceases to exist. One can generalize these models to include storage cost of physical money and allow for the interest rate to go below zero. In fact the Bank of Japan and the European Central Bank, among other central banks, have been able to reduce their policy rates below zero. Nonetheless, it remains plausible to assume that there is a bound beyond which it becomes very difficult to lower interest rates further and this what the literature considers to be the ELB.
problem, because the coefficient shift that takes place once the constrained becomes binding is controlled by a low-dimensional vector of additional parameters.

The paper makes the following contributions. First, we develop a novel SVAR model for a vector of observables \( y_t = [y_{1,t}, y_{2,t}'] \), where \( y_{1,t} \) is a censored dependent variable, \( y_{1,t} = \max\{y_{1,t}^*, 0\} \). We define the endogenous state variable \( s_t = \mathbb{1}\{y_{1,t} > 0\} \) and let the VAR coefficients that describe the law of motion of \( y_{2,t} \) depend on the state \( s_t \). In addition to parsimony, a second challenge in specifying a model of this form is to ensure that the outcome \( y_t \) is uniquely determined as a function of its lags and a set of innovations \( \epsilon_t \). This existence of a reduced form is referred to in the literature as coherency and completeness; see Mavroeidis (2020).

In every period \( t \) we can compute two hypothetical values for \( y_{1,t} \): one is based on the \( s_t = 1 \) coefficients and the other one is based on the \( s_t = 0 \) coefficients. For \( y_{1,t} \) to be uniquely determined, it is necessary that only one of these two values has the property that the indicator function \( \mathbb{1}\{y_{1,t} > 0\} \) coincides with the \( s_t \) that was used to compute \( y_{1,t} \). Typically, this uniqueness cannot be achieved without restricting the domain of the innovations \( \epsilon_t \).

Our key theoretical result is to prove that if the regression functions associated with \( y_{2,t} \) (in the context of a DSGE model they would represent the private-sector decision rules) are continuous in the regressors (the state variables in the context of a DSGE model) whenever the system reaches the censoring point, then under some mild restrictions on the VAR coefficients, uniqueness can be obtained for any value of regressors \( x_t \) and innovations \( \epsilon_t \).

Second, we use state-of-the-art Bayesian techniques to estimate our model. It is nonlinear and contains complicated restrictions across the coefficients of the \( s = 1 \) and \( s = 0 \) regime. Thus, the posterior distributions of the model parameters are non-standard. We show how to use a sequential Monte Carlo algorithm to implement Bayesian inference.

Third, our empirical analysis generates several novel findings. The model is estimated based on quarterly U.S. data from 1984:Q1 to 2018:Q4. The vector \( y_t \) in our empirical analysis comprises the federal funds rate as the censored variable, output gap, and inflation. We find evidence in favor of a kink in the inflation regression function. We disentangle two types of nonlinearities: the nonlinearity coming from the censoring and the nonlinearity from the kink in the regression function. The largest effect of censoring is visible in the response of the economy to a demand shock. It accelerates the reversion of output gap and inflation to their long-run means. The kink in the inflation equation amplifies the inflation response to
a monetary policy shock and switches the sign of the inflation response to a contractionary demand shock.

A comparison of impulse responses conditional on 1999:Q1 data (the U.S. was far away from the ELB) and conditional on 2009:Q1 data (the U.S. was at the ELB) shows significantly different responses of inflation to monetary policy and demand shocks with little overlap of the bands that represent posterior uncertainty. An unconventional expansionary monetary policy intervention at the ELB in 2009:Q1 is much more inflationary than away from the ELB in 1999:Q1. Moreover, the price adjustment in response to a demand shock is much more muted in 2009:Q1 and has the opposite sign than in 1999:Q1. Thus, we conclude that, as predicted by DSGE models with an ELB constraint, the ELB is not irrelevant for the propagation of shocks and this effect is measurable in a parsimonious SVAR framework that allows for changes in the private-sector behavior at the ELB. Our model also generates a shadow interest rate which dropped to -3% in 2009 and hovered around -1% between 2010 and 2015.

Our paper is related to several strands of the literature. From a methodological perspective, the paper most closely related to our work is Mavroeidis (2020). He also considers an SVAR with a censored dependent variable, or occasionally-binding constraint, to capture the ELB constraint on nominal interest rates. An important difference between his setup and our model specification is that in our specifications the coefficients of the private-sector equations are allowed to be state dependent, just as in a DSGE model.

Building on an older literature on simultaneous equations models with censored dependent variables, e.g., Nelson and Olsen (1978) and Blundell and Smith (1989), the emphasis in Mavroeidis (2020) is on the identifying information that the censoring provides for the propagation of structural shocks. Important for the identification is whether, in the case of the ELB application, the private-sector variables respond to the actual censored interest rate or the uncensored shadow interest rate. In our application, we let inflation and output gap respond to monetary policy shocks\(^2\), even if the economy is at the ELB, which is similar in spirit to the censored SVAR specification in Mavroeidis (2020) in which agents respond to a shadow rate. The identifying information encoded in the censoring mechanism is implicitly exploited in our Bayesian estimation through the updating of the structural VAR coefficients based on the likelihood function.

Abstracting from the nonlinearities generated by the censoring of the nominal interest

\(^2\)This is also the case in DGSE model solutions; see ACHSV
rates and the piecewise-linear regression equations for private-sector variables, the specification of the SVAR follows the three-variable model estimated in Baumeister and Hamilton (2018), henceforth BH. While the prior distribution for the VAR coefficients is not identical to the one used by BH, the elicitation of the prior distribution for the coefficients that describe the contemporaneous interaction between output, inflation, and interest rates follows a similar logic. As in BH, a linear version of our model that abstracts from the ELB constraint, would only be set identified. The prior combines beliefs about aggregate demand and supply elasticities formed based on a simple New Keynesian DSGE model with beliefs about directions of impulse responses.

To generate draws from the posterior distribution we use a sequential Monte Carlo (SMC) algorithm. The algorithm has been developed in the statistics literature and in recent years it has been applied to posterior inference in a variety of econometric time series models. Our implementation is based on the algorithm in Herbst and Schorfheide (2014, 2015) which was tailored toward the estimation of a DSGE model. Bognanni and Herbst (2018) applied a similar algorithm to the Bayesian estimation of an SVAR with exogenous regime switches that follow a Markov process.

In terms of the empirical application our paper is closely related to Debortoli, Galí, and Gambetti (2019) who estimate a time-varying coefficient SVAR to examine whether the ELB altered the effectiveness of monetary policy interventions and the propagation of shocks. While their empirical model is nonlinear, the nonlinearity is generated through exogenously time-varying coefficients (TVC). To avoid complications arising from the ELB constraint, they estimate their TVC-SVAR based on a longer-term interest rate that does not reach the ELB.

In our model, on the other hand, the nonlinearity is generated through the censoring of the nominal interest rate and the kink in the regression function that describes the private-sector behavior. A key challenge for TVC-SVARs is to generate precise estimates for a high-dimensional parameter vector over a short window of time. While our model allows for time-variation in coefficients, it does so in a much more parsimonious way. In a three-variable system we essentially use two parameters to capture changes in the private-sector regression functions. While Debortoli, Galí, and Gambetti (2019) find little evidence against the hypothesis that the ELB is irrelevant, our more parsimonious specification allows us to detect a significant change in coefficients that alters the propagation of monetary policy and demand shocks and lead us to conclude the ELB constraint is relevant for the propagation of shocks.
The remainder of the paper is organized as follows. The specification of our SVAR with a censored dependent-variable and state-dependent regression functions is presented in Section 2. Our prior distribution for the SVAR parameters is discussed in Section 3. The likelihood function is derived in Section 4 and the SMC algorithm to implement the posterior computations is summarized in Section 5. Section 6 presents the empirical analysis and Section 7 concludes. Theoretical derivations and additional empirical results are relegated to the Online Appendix.

2 SVAR Specification

We are using a structural vector autoregression (SVAR) to model the law of motion of the \( n \times 1 \) vector \( y_t \). There are two common ways of normalizing the coefficients of an SVAR: by setting the matrix of coefficients associated with the time \( t \) endogenous variables \( y_t \) equal to the identity matrix, or by setting the matrix that governs the impact of the structural shocks \( \epsilon_t \) equal to the identity matrix. We begin with the former normalization and call it the \( \Phi \) representation:

\[
y_t = \Phi_1 y_{t-1} + \ldots + \Phi_p y_{t-p} + \Phi_c + \Phi_\epsilon \epsilon_t,
\]

where \( \epsilon_t \sim N(0, D) \) is a vector of structural innovations and \( D \) is a diagonal matrix. Reduced-form innovations can be defined as

\[
u_t = \Phi_\epsilon \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma) \quad \text{where} \quad \Sigma = \Phi_\epsilon D \Phi_\epsilon'.
\]

Define the \( 1 \times k \) vector \( x_t' = [y_{t-1}', \ldots, y_{t-p}', 1]' \), \( \Phi = [\Phi_1, \ldots, \Phi_p, \Phi_c]' \), and \( \Phi^\epsilon = \Phi_\epsilon' \) so that we can write the VAR as

\[
y_t' = x_t' \Phi + \epsilon_t' \Phi^\epsilon.
\]

Let \( A = (\Phi^\epsilon)^{-1} \) and multiply (2) by \( A \) to re-normalize the VAR and obtain what we call the \( A \) representation of the VAR:

\[
y_t'A = x_t'B + \epsilon_t',
\]

where \( B = \Phi A \).

Starting point for the subsequent analysis will be a third representation, which we refer to as \( A\Phi \) representation. It combines a monetary policy rule written in \( A \) form with the private-sector equations in \( \Phi \) form and has been used, for instance, in Del Negro and Schorfheide (2009). The \( \Phi \)-form of the private-sector equations resemble the decision rules in a DSGE
model solution that are allowed to differ at and away from the ELB. This representation also facilitates our proof of the existence of a unique reduced form; see Proposition 1 in Section 2.3 below.

Partition $y_t' = [y_{1,t}, y_{2,t}]$ and $\epsilon_t' = [\epsilon_{1,t}, \epsilon_{2,t}]$, where $y_{1,t}$ corresponds to the interest rate and $\epsilon_{1,t}$ is the monetary policy shock. Moreover, partition $B = [B_1, B_2]$, where $B_1$ is a column vector that stacks the coefficients of the $y_{1,t}$ equation and the columns of the matrix $B_2$ stack the coefficients for the private-sector equations. Finally, partition $A_1 = [A_{11} | A_{21}]$, where we use $|$ to indicate that the partitions are stacked. Using this notation, the monetary policy rule becomes

$$y_{1,t} A_{11} + y_{2,t} A_{21} = x_t' B_1 + \epsilon_{1,t}. \tag{4}$$

The private-sector behavior is described in $\Phi$ form:

$$y_{2,t}' = x_t' \Phi_2 + \epsilon_{1,t} \Phi_{12} + \epsilon_{2,t}' \Phi_{22}. \tag{5}$$

In the remainder of this section, we will introduce two important extensions of the specification in (4) and (5). To capture the ELB constraint on nominal interest rates we allow for censoring of $y_{1,t}$ in Section 2.1. Second, in Section 2.2 we allow the private-sector behavior to change when the ELB becomes binding and examine the consequence of this generalization on the coherency of the model. Third, in Section 2.3 we analyze the special case of private-sector regression functions that remain continuous in the state variables when the ELB constraint becomes binding.

### 2.1 Censoring

In order to capture the ELB constraint, we introduce censoring. We use $y_{1,t}^*$ to denote the desired or shadow interest rate and write the monetary policy rule as

$$y_{1,t}^* A_{11} + y_{2,t} A_{21} = x_t' B_1 + \epsilon_{1,t}. \tag{6}$$

Here we replaced $y_{1,t}$ in (4) by $y_{1,t}^*$. The relationship between $y_{1,t}$ and $y_{1,t}^*$ is given by

$$y_{1,t} = \max \{y_{1,t}^*, 0\}. \tag{7}$$

We will assume that both central bank and the private sector react to lagged $y_{1,t}$ instead of $y_{1,t}^*$. However, we do allow agents to react to the monetary policy shock when the ELB is
binding. Both of these assumptions are consistent with the DSGE model in ACHSV. Thus, (5) remains unchanged. We define \( y^*_{t'} = [y^*_{1,t}, y^*_{2,t}] \).

### 2.2 Regime-Dependent Private-Sector Behavior

We will now extend the model and allow the private sector to change its behavior once the economy reaches the ELB. We introduce the regime (or ELB) indicator

\[
s_t = \mathbb{I}\{y^*_{1,t} > 0\}
\]

and write the private-sector equations as

\[
y'_{2,t} = x_t^s \Phi_2(s_t) + \epsilon_{1,t}^s \Phi_{12}^s(s_t) + \epsilon_{2,t}^s \Phi_{22}^s(s_t) = x_t^s \Phi_2(s_t) + u_{2,t}^s(s_t).
\]

Plugging the expression for \( y'_{2,t} \) from (9) into the monetary policy rule (6) leads to the \( \Phi \) form of the interest rate equation:

\[
y^*_{1,t} = \frac{1}{A_{11}} \left[ x_t^s (B_1 - \Phi_2(s_t)A_{21}) + \epsilon_{1,t}^s (1 - \Phi_{12}^s(s_t)A_{21}) - \epsilon_{2,t}^s \Phi_{22}^s(s_t)A_{21} \right].
\]

Define

\[
\Phi_1(s_t) = \frac{1}{A_{11}} (B_1 - \Phi_2(s_t)A_{21}),
\]

\[
u_{1,t}(s_t) = \frac{1}{A_{11}} \left[ \epsilon_{1,t}^s (1 - \Phi_{12}^s(s_t)A_{21}) - \epsilon_{2,t}^s \Phi_{22}^s(s_t)A_{21} \right]
\]

such that we can write

\[
y^*_{1,t} = x_t^s \Phi_1(s_t) + u_{1,t}(s_t).
\]

In view of (8), for the model specification to be internally consistent, it has to be the case that when the “1” regression functions are active that \( y_{1,t} > 0 \). Likewise, whenever the “0” regression functions are active it has to be the case that \( y^*_{1,t} \leq 0 \). Given a set of parameters \( A_1, B_1, \Phi_2(s), \Phi_2^s(s) \), lagged values \( x_t \), and a vector of structural shocks \( \epsilon_t \) we can distinguish three cases:

**Case 1 – Uniqueness:** conditional on lagged values \( x_t \) and the innovation \( \epsilon_t \) the state \( s_t \),
the latent variable $y_{1,t}^*$ and $y_{2,t}$ are uniquely determined. If

$$x_t'\Phi_{1}(1) + u_{1,t}(1) > 0, \quad \text{then} \quad x_t'\Phi_{1}(0) + u_{1,t}(0) > 0 \quad \text{which implies} \quad s_t = 1.$$  

Alternatively, if

$$x_t'\Phi_{1}(1) + u_{1,t}(1) \leq 0, \quad \text{then} \quad x_t'\Phi_{1}(0) + u_{1,t}(0) \leq 0 \quad \text{which implies} \quad s_t = 0.$$  

We use $E_t^U$ to denote the set of $\epsilon_t$ values for which $s_t$ is unique.

**Case 2 - Indeterminacy (Incompleteness):** conditional on lagged values $x_t$ and the innovation $\epsilon_t$, the model is consistent with $s_t = 0$ and $s_t = 1$, which means that there are two possible values for $y_{1,t}^*$ and $y_{2,t}$, respectively. Formally,

$$x_t'\Phi_{1}(1) + u_{1,t}(1) > 0 \quad \text{and} \quad x_t'\Phi_{1}(0) + u_{1,t}(0) \leq 0 \quad \text{which implies} \quad s_t = 1 \text{ or } s_t = 0.$$  

We use $E_t^I$ to denote the set of $\epsilon_t$ values for which $s_t$ is not unique.

**Case 3 - Non-existence (Incoherency):** conditional on lagged values $x_t$ and the innovation $\epsilon_t$, the model is neither consistent with $s_t = 0$ nor $s_t = 1$ because

$$x_t'\Phi_{1}(1) + u_{1,t}(1) \leq 0 \quad \text{and} \quad x_t'\Phi_{1}(0) + u_{1,t}(0) > 0.$$  

We use $E_t^N$ to denote the set of $\epsilon_t$ values for we have non-existence.

The fact that the existence and uniqueness of $y_t$ depend on the lagged endogenous variables stacked in $x_t$ and the structural innovations $\epsilon_t$ is an undesirable feature of the model. To rule out non-existence in general one would need to restrict the domain of the innovations $\epsilon_t$; see Mavroeidis (2019, 2020). We will show in the following section that in the special case of piecewise-linear regression functions, the uniqueness condition can be expressed as a restriction on the parameter space and does not depend on $(x_t, \epsilon_t)$.

### 2.3 Piecewise Linear and Continuous Regression Functions

Building on ACHSV, we now impose that the private sector uses regression functions that are continuous at the kink. We refer to these regression functions as piecewise linear and continuous (PLC). Consider the monetary policy rule in $\Phi$ form. From (10) we deduce that
the ELB starts to bind whenever

\[ x'_t(B_1 - \Phi_2(1)A_{21}) = -\epsilon_{1,t}(1 - \Phi_{12}(1)A_{21}) + \epsilon'_{2,t}\Phi'_{22}(1)A_{21}. \]  

(13)

Now partition \( x_t = [x'_{1,t}, x'_{2,t}] \), and, conformingly, partition \( B_1 = [B_{11}|B_{21}] \) and \( \Phi_2 = [\Phi_{12}|-\Phi_{22}] \). Using this notation, we can write (13) as

\[ x_{1,t} = -x'_{2,t}\frac{B_{21} - \Phi_{22}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} - \epsilon_{1,t}\frac{1 - \Phi_{12}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} + \epsilon'_{2,t}\frac{\Phi_{22}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}}. \]  

(14)

Continuity at the kink implies that

\[ x_{1,t}\Phi_{12}(1) + x'_{2,t}\Phi_{22}(1) + \epsilon_{1,t}\Phi'_{12}(1) + \epsilon'_{2,t}\Phi'_{22}(1) \]

\[ = x_{1,t}\Phi_{12}(0) + x'_{2,t}\Phi_{22}(0) + \epsilon_{1,t}\Phi'_{12}(0) + \epsilon'_{2,t}\Phi'_{22}(0). \]  

(15)

Now plug the expression for \( x_{1,t} \) in (14) into Equation (15) and use the continuity restrictions to solve for the coefficients in the \( s = 0 \) regime:

\[ \Phi_{22}(0) = \Phi_{22}(1) + \frac{B_{21} - \Phi_{22}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}}\Phi_{12}^\Delta, \]  

(16)

\[ \Phi'_{12}(0) = \Phi'_{12}(1) + \frac{1 - \Phi_{12}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}}\Phi_{12}^\Delta, \]

\[ \Phi'_{22}(0) = \Phi'_{22}(1) - \frac{\Phi_{22}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}}\Phi_{12}^\Delta, \]

where

\[ \Phi_{12}^\Delta = \Phi_{12}(0) - \Phi_{12}(1). \]

Here the unrestricted coefficient matrices are \( \Phi_{22}(1), \Phi'_{12}(1), \) and \( \Phi_{12}^\Delta \). Two special cases are noteworthy. First, if \( \Phi_{12}^\Delta = 0 \) then the regression functions are strictly linear and have no kink: \( \Phi_{22}(0) = \Phi_{22}(1) \) and \( \Phi'_{12}(0) = \Phi'_{12}(1) \). Second, if the private sector does not react to the monetary policy shock, i.e., \( \Phi_{12}(1) = 0 \) and \( \Phi'_{12}(0) = 0 \), then \( \Phi_{12}^\Delta = 0 \), which in turn implies that there is no kink in the reactions to \( x_{2,t} \) and \( \epsilon_{2,t} \): \( \Phi_{22}(0) = \Phi_{22}(1) \) and \( \Phi'_{22}(0) = \Phi'_{22}(1) \).

**Proposition 1 (Sufficient Condition)** If the regression functions are continuous at the ELB kink and one of the following two conditions is satisfied:

\( i \) \( B_{11} - \Phi_{12}(1)A_{21} > 0 \) and \( B_{11} - (\Phi_{12}(1) + \Phi_{12}^\Delta)A_{21} > 0, \)

\( ii \) \( B_{11} - \Phi_{12}(1)A_{21} < 0 \) and \( B_{11} - (\Phi_{12}(1) + \Phi_{12}^\Delta)A_{21} < 0, \)

then \( \mathbb{P}\{\epsilon_t \in \mathcal{E}_t^U\} = 1. \)
In our estimation, we will impose the parameter restrictions described in Proposition 1 to ensure uniqueness of \((s_t, y_{1,t}, y_{2,t})\) conditional on all values of \((x_t, \epsilon_t)\). After allowing for regime-dependent private-sector regression functions and imposing continuity at the kink, the parameters of the SVAR model are:

\[
\begin{align*}
\begin{array}{l}
A_1, \quad B_1, \quad \Phi_2(1), \quad \Phi_{\epsilon_1}(1), \quad \Phi_{\Delta 12}, \quad D.
\end{array}
\end{align*}
\]

3 Prior Distribution

We use Bayesian techniques to estimate the SVAR with ELB censoring and piecewise-linear regression functions. This requires the specification of a prior distribution. It is important to note that without further restrictions, and in particular if the censoring mechanism is not triggered, the SVAR is only set-identified. In set-identified models there always exist functions of the model parameters for which the prior distribution is not updated; see, for instance, Poirier (1998) and Moon and Schorfheide (2012).

In our empirical application we assume that \(y_{1,t}\) corresponds to the nominal interest rate \(R_t\), and \(y_{2,t} = [z_t, \pi_t]\), where \(z_t\) is output gap and \(\pi_t\) is inflation. We interpret the innovations \(\epsilon_t = [\epsilon_{R,t}, \epsilon_{D,t}, \epsilon_{S,t}]\) as monetary policy, demand, and supply shocks, respectively. Thus, in the \(A\) representation the first equation is the monetary policy rule and the remaining two equations are demand and supply. Our prior for the elements of the \(A(1)\) matrix builds on the prior used by Baumeister and Hamilton (2018). They provide an informal calculation that relates the demand equation to the Euler equation and the supply equation to the New Keynesian Phillips curve in a small-scale New Keynesian DSGE model.

3.1 (Re)parameterization of the SVAR

To facilitate the elicitation of a prior distribution, we replace the \(n \times (n-1)\) matrix \(\Phi_{\epsilon_2}(s)\) by the \(n \times (n-1)\) matrix \(A_2(s)\). We solve (4) for \(\epsilon_{1,t}\) and plug the expression into (5). This yields:

\[
y'_{2,t} = x'_t \Phi_2(s) + (y_{1,t} A_{11} + y'_{2,t} A_{21} - x'_t B_1) \Phi_{12}(s) + \epsilon'_{2,t} \Phi_{\epsilon_2}(s).
\]

After re-arranging terms we obtain

\[
-y_{1,t} A_{11} \Phi_{12}(s) + y'_{2,t} (I - A_{21} \Phi_{12}(s)) = x'_t (\Phi_2(s) - B_1 \Phi_{12}(s)) + \epsilon'_{2,t} \Phi_{\epsilon_2}(s). \quad (17)
\]
Let $A_2(s) = [A_{12}(s) | A_{22}(s)]$ and deduce that

$$A_{12}(s) = -A_{11} \Phi_{12}'(s)(\Phi_{22}'(s))^{-1}, \quad A_{22}(s) = (I - A_{21} \Phi_{12}'(s)) (\Phi_{22}'(s))^{-1}. \quad (18)$$

Now define $A(s) = [A_1, A_2(s)]$. We follow Baumeister and Hamilton (2018) by expressing the matrix of contemporaneous interactions, in our case for the non-binding regime $s = 1$, as a function of the parameters $[\rho, \psi_{\pi}, \psi_{z}, \alpha_{S}, \beta_{D}, \gamma_{D}]'$:

$$A(1) = \begin{bmatrix} 1 & -\gamma_{D} & 0 \\ -(1 - \rho) \psi_{z} & 1 & 1 \\ -(1 - \rho) \psi_{\pi} & -\beta_{D} & -\alpha_{S} \end{bmatrix} \quad (19)$$

Recall that the observables are ordered as follows: $y_t = [R_t, z_t, \pi_t]$. The parameterization is motivated by a three-equation New Keynesian DSGE model. The first column of $A(1)$ can be interpreted as a Taylor rule with interest rate smoothing $\rho$ and output and inflation gap coefficients $\psi_{z}$ and $\psi_{\pi}$, respectively.

The second column represent the Euler equation after the expectations of next period’s inflation and output gap have been replaced by AR(1) forecasts. This creates a contemporaneous relationship between output gap (being proportional to consumption), the nominal interest rate, and inflation. The parameter $\beta_{D}$ can be interpreted as an aggregate demand elasticity.

The third column represents the New Keynesian Phillips curve with marginal costs replaced by output gap and expected inflation replaced by an AR(1) forecast. The coefficient $\alpha_{S}$ can be interpreted as a supply elasticity. We partition $B_{1} = [B_{11} | B_{21}]$, where $B_{11}$ is scalar and represents the coefficient on $R_{t-1}$. We interpret this coefficient as the degree of interest-rate smoothing and set it equal to $B_{11} = \rho$.

The subsequent prior distribution is specified over the set of parameters

$$\theta = (\rho, \psi_{\pi}, \psi_{z}, \alpha_{S}^{\Delta}, \beta_{D}, \gamma_{D}, B_{21}, \Phi_{2}(1), \ldots, \Phi_{n}(1), \Phi_{12}^{\Delta}, D_{11}, \ldots, D_{nn}),$$

where $\alpha_{S}^{\Delta} = \alpha_{S} - \beta_{D}$. In slight abuse of notation, the vectors $\Phi_{i}(1)$ refer to the $n - 1$ columns of the matrix that we previously denoted as $\Phi_{2}(1)$ and the $D_{ii}$’s are the elements of the diagonal matrix $D$. It is important to note that despite the restrictions we imposed on $A(1)$,
the SVAR remains set-identified. In the absence of censoring and regime-dependent private-
sector regression functions, one would have to impose three equality restrictions on the $A$
matrix. Here we are only imposing one restriction in addition to the three normalizations,
namely, the zero in the upper right corner. As in Baumeister and Hamilton (2018), we will
add more information by restricting the signs of the elasticities $\alpha_S$, $\beta_D$, and $\gamma_D$.

3.2 Distributional Assumptions

Our prior distribution has a hierarchical structure with hyperparameters $(\phi, \lambda_d, \lambda_0, \lambda_l, \lambda_I, \lambda_\Delta)$ and takes the following form:

$$
\tilde{p}(\theta) = p(\rho, \psi_\pi, \psi_z, \alpha_\Delta^S, \beta_D, \gamma_D) \left( \prod_{i=1}^{n} p(D_{ii} | \vartheta; \lambda_d) \right) \times p(B_{21} | D, \rho; \lambda_0, \lambda_l, \lambda_I) \left( \prod_{i=2}^{n} p(\Phi_i(1) | \vartheta, D; \phi, \lambda_0, \lambda_l, \lambda_I) \right) p(\Phi_{12}^\Delta | \lambda_\Delta).
$$

(20)

Rather than including the full list of parameters $(\rho, \psi_\pi, \psi_z, \alpha_\Delta^S, \beta_D, \gamma_D)$ into the condition sets, we simply abbreviate them by $\vartheta = (\rho, \psi_\pi, \psi_z, \alpha_\Delta^S, \beta_D, \gamma_D)$.

**Prior for Contemporaneous Interactions** $p(\rho, \psi_\pi, \psi_z, \alpha_\Delta^S, \beta_D, \gamma_D)$. We assume that the underlying parameters are independent:

$$
p(\rho, \psi_\pi, \psi_z, \alpha_\Delta^S, \beta_D, \gamma_D) = p(\rho)p(\psi_\pi)p(\psi_z)p(\alpha_\Delta^S)p(\beta_D)p(\gamma_D).
$$

We impose the following domain/sign restrictions on the coefficients:

$$
0 \leq \rho < 1, \quad \psi_\pi \geq 0, \quad \psi_z \geq 0, \quad \alpha_\Delta^S \geq 0, \quad \beta_D \geq 0, \quad \gamma_D \leq 0.
$$

(21)

The restriction of $\beta_D \geq 0$ may seems counterintuitive in view of the interpretation of the coefficient as a demand elasticity. However, it is consistent with a DSGE model in which the demand equation represents a consumption Euler equation and a rise in (expected) inflation lowers real returns which creates an incentive to increase current-period consumption.

To facilitate the elicitation of a prior distribution, it is helpful to derive the inverse of
Table 1: Prior Distribution for $A(1)$ Coefficients

<table>
<thead>
<tr>
<th>Param.</th>
<th>Distr.</th>
<th>P(1)</th>
<th>P(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>Beta</td>
<td>0.8</td>
<td>0.1</td>
</tr>
<tr>
<td>$\psi_\pi$</td>
<td>Gamma</td>
<td>1.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$\psi_z$</td>
<td>Gamma</td>
<td>0.5</td>
<td>0.25</td>
</tr>
<tr>
<td>$\alpha^\Delta S$</td>
<td>Gamma</td>
<td>1.0</td>
<td>0.7</td>
</tr>
<tr>
<td>$\beta_D$</td>
<td>Gamma</td>
<td>1.0</td>
<td>0.7</td>
</tr>
<tr>
<td>$-\gamma_D$</td>
<td>Gamma</td>
<td>0.25</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Notes: The table lists marginal distributions. P(1) and P(2) are means and standard deviations for Beta and Gamma distributions.

$A(1)$, which determines the contemporaneous impact of the three innovations:

$$[A(1)]^{-1} \propto \begin{bmatrix} \alpha_S - \beta_D & \alpha_S \gamma_D & \gamma_D \\ \zeta_\pi + \alpha_S \zeta_z & \alpha_S & 1 \\ -(\zeta_\pi + \beta_D \zeta_z) & -(\beta_D + \gamma_D \zeta_\pi) & \gamma_D \zeta_z - 1 \end{bmatrix}. \quad (22)$$

Here $\propto$ denotes proportionality. The first row of the matrix captures the response of interest rates, the output gap, and inflation to a monetary policy shock $\epsilon_{1,t}$. In order for interest rates to rise, and output and inflation to fall in response to a contractionary monetary policy shock we need $\alpha^\Delta S \geq 0$, $(\alpha^\Delta S + \beta_D) \geq 0$, and $\gamma_D \leq 0$ which is implied by (21).

As long as $\alpha_S$ is positive, output gap and inflation move in the same direction in response to a monetary policy shock. The parameter $\alpha_S$ controls the relative response of output gap and inflation and the ratio $\gamma_D/(\alpha_S - \beta_D)$ is the relative response of interest rates and inflation. According to the second row of $[A(1)]^{-1}$, output gap and prices move in the same direction provided $\alpha_S \geq 0$. Finally, the third row of the matrix in (22) determines the response to a demand shock. The direction of the inflation and output gap response to a supply shock are depend on the specific parameterization. For instance, assuming that $\gamma_D < 0$, inflation is falling in response to the supply shock. The output response is given by $-(\beta_D + \zeta_\pi \gamma_D)$. We observe a rise in output if $-\gamma_D > \beta_D/\zeta_\pi$.

Table 1 summarizes the specification of the benchmark prior for the $A(1)$ coefficients. The priors for $(\rho, \psi_\pi, \psi_z)$ are broadly in line with priors used in the DSGE model literature. The numerical values for $(\alpha^\Delta S, \beta_D, -\gamma_D)$ are more difficult to interpret. In the empirical section we will plot the prior distribution of impulse responses.

**Prior for Innovation Variances** $p(D_{ui}|\theta; \lambda_d)$. We fit univariate AR(p) models to the series
\(y_{i,t}\) and let \(s_i^2\) be the estimated innovation variance. We use \(\vartheta\) to generate the \(A(1)\) matrix in (19) and define the matrix

\[
S = A(1)' \begin{bmatrix}
s_1^2 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & s_n^2
\end{bmatrix} A(1).
\]

We specify a prior for \(\sqrt{D_{ii}}, i = 1, \ldots, n\). Starting point is a prior for the \(D_{ii}\) elements, which takes the form of an Inverse Gamma distribution that is parameterized as scaled Inverse \(\chi^2\) distribution with density

\[
p(D_{ii}|\sqrt{S_{ii}}, \lambda) = \left(\frac{\lambda}{2}\right)^{\lambda_d/2} \left(\frac{S_{ii}}{\lambda}\right)^{\lambda_d/2} \Gamma\left(\frac{\lambda_d}{2}\right) D_{ii}^{-\lambda_d/2-1} \exp\left\{-(\lambda_d/2)S_{ii}D_{ii}^{-1}\right\}.\]

The density of \(\sqrt{D_{ii}}\) is obtained by a change-of-variable, which adds the Jacobian term \(2\sqrt{D_{ii}}\) to the density. Here, \(\lambda_d\) is a hyperparameter that controls the degrees of freedom of the scaled Inverse \(\chi^2\) distribution.

**Prior for Policy Rule Coefficients** \(p(B_{21}|D, \rho; \lambda_0, \lambda_l, \lambda_I)\). We will now specify a prior for \(B_{21}\). Note that the innovation variance for the interest rate equation is simply given by \(D_{11}\). Recall that \(x_t' = [y_{1,t-1}, y_{2,t-1}, y_{t-2}, \ldots, y_{t-p}, 1]\). Because \(y_{1,t-1} = R_{t-1}\), we set \(B_{11} = \rho\). Recall that we denoted the remaining elements by \(B_{21}\), which is a column vector with \(k - 1\) elements. Let

\[
B_{21}|(D, \lambda_0, \lambda_l, \lambda_I) \sim N(B_{21}, P_{21}^{-1}(\lambda_0, \lambda_l, \lambda_I)).
\]

We set the prior mean \(B_{21} = 0\). The \((k - 1) \times (k - 1)\) precision matrix \(P_{21}(\lambda_0, \lambda_l, \lambda_I)\) is diagonal with the following elements: (i) the value associated with the coefficient of the \(l\)th lag of variable \(j\) is \(l\lambda s_j^2/(\lambda D_{11})\), where \(s_j\) was defined above. (ii) The value associated with the intercept is set to \(1/(\lambda_I D_{11})\). Here \(\lambda_0\) is a hyperparameter that scales the overall variance of the prior; \(\lambda_l\) determines how quickly the prior variance decays with lag length \(l\), and \(\lambda_I\) is the scale of the prior variance for the intercept.

**Prior for Private-Sector Coefficients** \(p(\Phi_{i}(1)|\vartheta, D; \phi, \lambda_0, \lambda_l, \lambda_I)\). We will now specify a prior for the \(k \times 1\) column vectors \(\Phi_{i}(1)\) that stack the reduced-form coefficients for private-sector variable \(i\), where \(i = 2, \ldots, n\). From (3) we deduce that the reduced-form forecast errors are given by

\[
u_t' = \epsilon_t'[A(1)]^{-1},
\]
as before, the $A(1)$ matrix is generated from $\vartheta$. The covariance matrix of the forecast errors is
\[
\Sigma(A, D) = [A(1)]^{-1'}E[\epsilon_t \epsilon_t'] [A(1)]^{-1} = [A(1)]^{-1'}D[A(1)]^{-1}.
\]

Let
\[
\Phi_{i}(1) \mid (D; \phi, \lambda_0, \lambda_l, \lambda_I) \sim N(\Phi_i(\phi), P_i^{-1}(\lambda_0, \lambda_l, \lambda_I)), \quad i = 2, \ldots, n. \tag{24}
\]

In order to impose the belief that the individual series are well approximated by AR(1) processes we set the prior mean vector as follows: $\Phi_{ii}(\phi) = \phi$ — this coefficient interacts with $y_{t-1}$ and $\Phi_{ji}(\phi) = 0$ for $j \neq i$. The $k \times k$ precision matrix $P_i(\lambda_0, \lambda_l, \lambda_I)$ is diagonal with the following elements: (i) the value associated with the coefficient of the $l$th lag of variable $j$ is $l^2 \lambda_I s_j^2 / (\lambda_0 \Sigma_{ii})$, where $s_j$ was defined above. (ii) The value associated with the intercept is set to $1 / (\lambda_I \Sigma_{ii})$. As before, $\lambda_0$ is a hyperparameter that scales the overall variance of the prior; $\lambda_l$ determines how quickly the prior variance decays with lag length $l$, and $\lambda_I$ is the scale of the prior variance for the intercept.

**Prior for Regression Function Differentials** $p(\Phi^\Delta_{12})$. We assume that
\[
\Phi^\Delta_{12} | \lambda_\Delta \sim N(0, \lambda_\Delta I). \tag{25}
\]

Note that for $\Phi^\Delta_{12} = 0$ the regression functions remain unchanged once the economy reaches the ELB.

**Hyperparameter Choices.** While in principle we could choose hyperparameters that maximize the marginal data density, we simply fix them a priori. We set
\[
\phi = 0.75, \quad \lambda_d = 4, \quad \lambda_0 = 0.1, \quad \lambda_l = 1.5, \quad \lambda_I = 3, \quad \lambda_\Delta = 0.25.
\]

Similar choices are used in Baumeister and Hamilton (2018) and are based on the experience with Minnesota Priors; see Doan, Litterman, and Sims (1984).

### 3.3 Adjustments to the Baseline Prior

We make three adjustments to the prior $\tilde{p}(\theta)$ in (20). These adjustments are implemented as follows:
\[
p(\theta) \propto \tilde{p}(\theta) f_u(\theta) f_s(\theta) f_m(\theta). \tag{26}
\]
Uniqueness. We enforce the uniqueness restriction in Proposition 1. Define $D(\theta) = B_{11} - \Phi_{12}(1)A_{21}$, with the understanding that here $\Phi_{12}$ refers to the first row of the $k \times (n - 1)$ matrix $\Phi_2$. Then, with slight abuse of notation, let

$$f_u(\theta) = \mathbb{I}\{D(\theta) > 0\} \mathbb{I}\{\Phi_{12}^\Delta A_{21} > -D(\theta)\} + \mathbb{I}\{D(\theta) < 0\} \mathbb{I}\{\Phi_{12}^\Delta A_{21} > D(\theta)\}.$$  

Stationarity. We impose that the reduced form representation of the SVAR is stationary conditional on $s_t = 1$ and $s_t = 0$ for all $t$, respectively. To do so, we convert the $A\Phi$ representation of the SVAR into the $\Phi$ representation in (1) and check that all roots of the characteristic polynomial $I - \sum_{l=1}^{p} \Phi_l(s)z^l$ are outside of the unit circle for $s = 1$ and $s = 0$. Let $f_s(\theta)$ be the indicator function that is equal to one of the stationarity condition is satisfied and equal to zero otherwise.

Long-run Mean. Following Sims and Zha (1998), we use dummy observations $(y_*, x_*)$ to impose the belief that if all lagged variables are equal to their sample means, then the current variable will also be equal to its long-run mean. Let $\bar{y}$ denote the sample mean and define $y_* = \lambda_* \bar{y}$ and $x_* = [\lambda_* \bar{y}, \ldots, \lambda_* \bar{y}, \lambda_*]$. Then, define $f_m(\theta) = p(y_*|x_*, \Phi(1), \Sigma(1))$, where $\Phi(1)$ and $\Sigma(1)$ are the coefficient matrices for the reduced-form $\Phi$ representation of the SVAR, conditional on $s = 1$. Note that this adjustment function generates an $a$ priori correlation between the intercept and the coefficients on the lags conditional on $s = 1$. In our application we set $\lambda_* = 10$.

4 Likelihood Function

We now derive the likelihood associated with the SVAR model discussed in Section 2 under the assumption that the uniqueness condition in Proposition 1 is satisfied. We factorize the likelihood function as follows

$$p(Y_{1:T}|\theta) = \prod_{t=1}^{T} p(y_{1,t}|Y_{1:t-1}, \theta)p(y_{2,t}|y_{1,t}, Y_{1:t-1}, \theta), \quad (27)$$

where $Y_{1:t_2}$ denotes the sequence $y_{t_1}, \ldots, y_{t_2}$.

Parameter Transformations. We begin with several parameter transformations to obtain the $\Phi$ representation of the SVAR. Based on $\theta$ we can compute

$$A_{1,*}, A_{2}(1), B_{1,*}, \Phi_{2}(1), \Phi_{12}^\Delta, D.$$
From \( A(1) \) we obtain \( \Phi^\epsilon(1) = [A(1)]^{-1} \). This leads to the \( \Phi \) form for \( y_{2,t} \) when \( s_t = 1 \):

\[
y'_{2,t} = x'_t \Phi_{2,2}(1) + \epsilon_{1,t} \Phi_{12}(1) + \epsilon'_{2,t} \Phi_{22}(1).
\]

(28)

We proceed by transforming the monetary policy rule so that we obtain the \( \Phi \) form for \( y_{1,t} \).

Plugging the expression for \( y_{2,t} \) in (28) into the monetary policy rule (4) we obtain:

\[
y_{1,t} = \frac{1}{A_{11}} \left[ x'_t (B_1 - \Phi_{2,1}A_{21}) + \epsilon_{1,t} (1 - \Phi_{12}(1)A_{21}) - \epsilon'_{2,t} \Phi_{22}(1)A_{21} \right]
\]

(29)

We deduce that

\[
\Phi_{1,1}(1) = \frac{1}{A_{11}} (B_1 - \Phi_{2,1}A_{21}), \quad \Sigma(1) = \Phi^\epsilon(1)D\Phi^\epsilon(1).
\]

(30)

To obtain the \( \Phi \) form for \( s_t = 0 \), let \( \Phi_{12}(0) = \Phi_{12}(1) + \Phi_{12}' \) and use (16) to compute \( \Phi_{2}(0) \) and \( \Phi_{2}(0) \). Then follow the steps in (28) to (30) to obtain \( \Phi(0) \) and \( \Sigma(0) \).

**Period-\( t \) Densities.** We partition the covariance matrix \( \Sigma(s) \) into \( [\Sigma_{ij}(s)] \) such that the partitions conform with \( y_t = [y_{1,t}, y'_{2,t}]' \) and define

\[
\Sigma_{2|1}(s) = \Sigma_{22}(s) - \Sigma_{21}(s) \Sigma_{11}^{-1}(s) \Sigma_{12}(s), \quad M_{12}(s) = \Sigma_{11}^{-1}(s) \Sigma_{12}(s), \quad M_{21}(s) = M'_{12}(s).
\]

Moreover, let

\[
u_{1,t}(s_t) = y_{1,t} - x'_t \Phi_{1,1}(s_t) \quad u'_{2,t}(s_t) = y'_{2,t} - x'_t \Phi_{2,2}(s_t).
\]

The density \( p(y_{1,t}|\cdot) \) is obtained as follows. If the conditions in Proposition 1 are satisfied, then \( y_{1,t} = 0 \) whenever \( u_{1,t}(1) \leq -x'_t \Phi_{1,1}(1) \). Thus, the distribution of \( y_{1,t} \) is a mixture of a pointmass at zero and a continuous distribution with support on \((0, \infty)\). We write its density as

\[
p(y_{1,t}|\cdot) = \mathbb{I}\{y_{1,t} = 0\} F_N \left( -\frac{x'_t \Phi_{1,1}(1)}{\sqrt{\Sigma_{11}(1)}} \right) + \mathbb{I}\{y_{1,t} > 0\} p_N(u_{1,t}(1); \Sigma_{11}(1)).
\]

(31)

The derivation of \( p(y_{2,t}|y_{1,t},\cdot) \) is more tedious and relegated to the Online Appendix. We have to distinguish between \( y_{1,t} = 0 \) and \( y_{1,t} > 0 \). The densities are given by

\[
p(y_{2,t}|y_{1,t} > 0, \cdot) = p_N(u_{2,t}(1); u_{1,t} M_{12}(1), \Sigma_{2|1}(1)).
\]

(32)
and
\[ p(y_{2,t}|y_{1,t} = 0, \cdot) \]  
\[ = (2\pi)^{-(n-1)/2}|\Sigma_{21}(0)|^{-1/2}|\Sigma_{11}(0)|^{-1/2}|\tilde{V}_u(0)|^{1/2} \left[ F_N \left( -\frac{x_2\Phi_1(1)}{\sqrt{\Sigma_{11}(1)}} \right) \right]^{-1} \times \frac{\exp \left\{ \frac{1}{2} u_{2,t}^\prime(0)\Sigma_{21}^{-1}(0)u_{2,t}(0) + \frac{1}{2}\tilde{V}_u^{-1}(0)\tilde{M}_u u_{2,t}(0) \right\}}{\tilde{V}_u} 
\]  
\[ \times F_N \left( -\frac{x_2\Phi_1(0) + \tilde{M}_u u_{2,t}(0)}{\sqrt{\tilde{V}_u(0)}} \right), \]

where
\[ \tilde{M}_u = \frac{M_{12}\Sigma_{21}^{-1}}{M_{12}\Sigma_{21}^{-1}M_{21} + \Sigma_{11}^{-1}}, \quad \tilde{V}_u = \left( M_{12}\Sigma_{21}^{-1}M_{21} + \Sigma_{11}^{-1} \right)^{-1}. \]

Expressions (31), (33), and (32) can be plugged into (27) to evaluate the likelihood function.

5 Posterior Computations via SMC

Because of fairly complicated nonlinear parameter restrictions generated by the piecewise-linear structure of our SVAR, the posterior distribution of the parameters is non-standard. We use a sequential Monte Carlo (SMC) sampler to generate draws from the posterior of \( \theta \). SMC techniques have emerged as an attractive alternative to MCMC methods. SMC algorithms can be easily parallelized and, properly tuned, may produce more accurate approximations of posterior distributions than MCMC algorithms. Chopin (2002) showed how to adapt particle filtering techniques to conduct posterior inference for a static parameter vector. Textbook treatments of SMC algorithms are provided, for instance, by Liu (2001) and Cappé, Moulines, and Ryden (2005). This section provides a brief description of the algorithm used subsequently. It draws heavily from the more detailed exposition Herbst and Schorfheide (2014, 2015).

SMC combines features of classic importance sampling and modern MCMC techniques. The starting point is the creation of a sequence of intermediate or bridge distributions \( \{\pi_n(\theta)\}_{n=0}^{N_P} \) that converge to the target posterior distribution, i.e., \( \pi_{N_P}(\theta) = \pi(\theta) \). At any stage the (intermediate) posterior distribution \( \pi_n(\theta) \) is represented by a swarm of particles...
\{\theta_n^i, W_n^i\}_{i=1}^N \) in the sense that the Monte Carlo average

\[
\bar{h}_{n,N} = \frac{1}{N} \sum_{i=1}^N W_n^i h(\theta^i) \xrightarrow{a.s.} \mathbb{E}_{\pi_n}[h(\theta_n)]
\] (34)

as \( N \to \infty \), for each \( n = 0, \ldots, N_\phi \). We adopt the convention that the weights \( W_n^i \) are normalized to average to one. The bridge distributions are posterior distributions constructed from stage-\( n \) likelihood functions:

\[
\pi_n(\theta) = \frac{p_n(Y|\theta)p(\theta)}{\int p_n(Y|\theta)p(\theta)d\theta}
\] (35)

with the convention that \( p_0(Y|\theta) = 1 \), i.e., the initial particles are drawn from the prior, and \( p_{N_\phi}(Y|\theta) = p(Y|\theta) \). We use likelihood tempering of the form

\[
p_n(Y|\theta) = [p(Y|\theta)]^{\phi_n}, \quad \phi_n = \left( \frac{n}{N_\phi} \right)^\ell.
\] (36)

The tuning parameter \( \ell \) controls the shape of the tempering schedule.

The SMC algorithm proceeds iteratively from \( n = 0 \) to \( n = N_\phi \). Starting from stage \( n-1 \) particles \( \{\theta_{n-1}^i, W_{n-1}^i\}_{i=1}^N \) each stage \( n \) of the algorithm targets the posterior \( \pi_n \) and consists of three steps: correction, that is, reweighting the stage \( n-1 \) particles to reflect the density in iteration \( n \); selection, that is, eliminating a highly uneven distribution of particle weights (degeneracy) by resampling the particles; and mutation, that is, propagating the particles forward using a Markov transition kernel to adapt the particle values to the stage \( n \) bridge density.

**Algorithm 1 (Generic SMC Algorithm)**

1. **Initialization.** (\( n = 0 \) and \( \phi_0 = 0 \).) Draw the initial particles from the prior: \( \theta_1^i \overset{\text{iid}}{\sim} p(\theta) \) and \( W_1^i = 1, \ i = 1, \ldots, N \).

2. **Recursion.** For \( n = 1, \ldots, N_\phi \),

   (a) **Correction.** Reweight the particles from stage \( n-1 \) by defining the incremental weights

   \[
   \bar{w}_n^i = \frac{p_n(Y|\theta_{n-1}^i)}{p_{n-1}(Y|\theta_{n-1}^i)}
   \] (37)
and the normalized weights
\[
\bar{W}_i^n = \frac{\bar{w}_i^n W_i^{n-1}}{N \sum_{i=1}^{N} \bar{w}_i^n W_i^{n-1}}, \quad i = 1, \ldots, N.
\] (38)

(b) Selection (Optional). Resample the swarm of particles, \(\{\theta_i^{n-1}, \bar{W}_i^{n-1}\}_{i=1}^{N}\), and denote resampled particles by \(\{\hat{\theta}_i^n, W_i^n\}_{i=1}^{N}\), where \(W_i^n = 1\) for all \(i\).

(c) Mutation. Starting from \(\hat{\theta}_i^n\), propagate the particles \(\{\hat{\theta}_i^n, W_i^n\}\) via \(N_{MH}\) steps of a Metropolis-Hastings (MH) algorithm with transition density \(K_n(\theta|\bar{\theta}; \zeta_n)\) and stationary distribution \(\pi_n(\theta)\). Note that the weights are unchanged, and denote the mutated particles by \(\{\theta_i^n, W_i^n\}_{i=1}^{N}\).

3. The importance sampling approximation of the posterior mean \(\mathbb{E}_n[h(\theta)]\) is given by:
\[
\bar{h}_{N, \phi, N} = \sum_{i=1}^{N} h(\theta_i^N \phi) W_i^N \phi.
\] (39)

Moreover, the marginal likelihood can be approximate by
\[
\hat{p}(Y) = \prod_{n=1}^{N_{\phi}} \left( \frac{1}{N} \sum_{i=1}^{N} \bar{w}_i^n W_i^{n-1} \right).
\] (40)

The correction step is a classic importance sampling step, in which the particle weights are updated to reflect the stage \(n\) distribution \(\pi_n(\theta)\). The selection step is optional. On the one hand, resampling adds noise to the Monte Carlo approximation, which is undesirable. On the other hand, it equalizes the particle weights, which increases the accuracy of subsequent importance sampling approximations. The decision of whether or not to resample can be based on a threshold rule for the variance of the particle weights which can be transformed into an effective particle sample size:
\[
\hat{ESS}_n = N / \left( \frac{1}{N} \sum_{i=1}^{N} (\bar{W}_n^i)^2 \right).
\] (41)

If the particles have equal weights, then \(\hat{ESS}_n = N\). If one particle has weight \(N\) and all other particles have weight \(0\), then \(\hat{ESS}_n = 1\). To balance the trade-off between adding noise and equalizing particle weights, the resampling step is typically executed if \(\hat{ESS}_n\) falls below a threshold \(\underline{N}\), e.g., \(N/2\) or \(N/3\). An overview of specific resampling schemes is provided, for
instance, in the books by Liu (2001) or Cappé, Moulines, and Ryden (2005) (and references cited therein). We are using systematic resampling in the applications below.

The mutation step changes the particle values. In the absence of the mutation step, the particle values would be restricted to the set of values drawn in the initial stage from the prior distribution. This would clearly be inefficient, because the prior distribution is typically a poor proposal distribution for the posterior in an importance sampling algorithm. As the algorithm cycles through the $N_\phi$ stages, the particle values successively adapt to the shape of the posterior distribution. This is the key difference between SMC and classic importance sampling. The transition kernel $K_n(\theta|\tilde{\theta},\zeta_n)$ is designed to have the following invariance property:

$$\pi_n(\theta_n) = \int K_n(\theta_n|\tilde{\theta}_n;\zeta_n)\pi_n(\tilde{\theta}_n)d\tilde{\theta}_n. \quad (42)$$

Thus, if $\tilde{\theta}_{n}^i$ is a draw from $\pi_n$, then so is $\theta_{n}^i$. The mutation step can be implemented by using one or more steps of a MH algorithm. The probability of mutating the particles can be increased by blocking the elements of the parameter vector $\theta$ or by iterating the MH algorithm over multiple steps. The vector $\zeta_n$ summarizes the tuning parameters of the MH algorithm.

The SMC algorithm can be initialized for $n = 0$ with draws from the prior density. Unfortunately, because of the adjustment functions $f_u(\theta)$, $f_s(\theta)$, and $f_m(\theta)$, it is not possible to directly sample from the prior density $p(\theta)$ in (26). Instead, we apply the SMC algorithm twice. Starting with draws from the initial prior $\bar{p}(\theta)$, which admits direct sampling, we first use the bridge distributions

$$p_n(\theta) \propto \bar{p}(\theta)\left[f_u(\theta)f_s(\theta)f_m(\theta)\right]^{\phi_n}.$$ 

The last stage of the first run will provide us with draws from $p(\theta)$. We then reset $\phi_n = 0$ and, in a second SMC run initialized with the $p(\theta)$ draws, generate approximations to the sequence of posteriors $\pi_n(\theta)$ in (35).

6 Empirical Analysis

We now estimate a three-variable SVAR on U.S. quarterly data. $y_{1,t}$ is defined as the federal funds rate and $y_{2,t}$ comprises measures of output gap and inflation. Output gap is defined as log real GDP minus the log potential output series published by the Congressional Budget
Notes: Grey bands indicate the ELB period from 2009Q1 to 2015Q4, during which the federal funds rate is below 25bp. Vertical lines denote dates for which we compute impulse response functions, which are 1999Q1 and 2009Q1.

Office. We take inflation to be the year-over-year changes (\(\ln P_t - \ln P_{t-4}\)) in the personal consumption expenditure deflator. The three series are plotted in Figure 1. Because, unlike in our model, the effective federal funds rate was never exactly equal to zero when the economy reached the ELB, we set interest rates below 25 basis points (bp) equal to zero. The ELB episode is indicated by the gray band in the data plot.

The subsequent empirical analysis is based on observations from 1984:Q1 to 2018:Q4. We
Figure 2: Priors and Posterior for Selected $A(1)$ Parameters

\[
\begin{align*}
\alpha_S & \\
\beta_D & \\
\gamma_D & \\
\gamma_D / (\alpha_S - \beta_D) &
\end{align*}
\]

Notes: Base prior is $\tilde{p}(\theta)$ and adjusted prior is $p(\theta)$; see (26).

use $p = 1$ lags for the SVAR. We use the SMC algorithm described in Section 5 to generate draws $\theta^i$, $i = 1, \ldots, N$ from the posterior distribution. All the results presented subsequently are based on transformations of these draws. The parameter estimates are discussed in Section 6.1, the estimated impulse response functions are presented in Section 6.2, and the implied shadow rate is discussed in Section 6.3.

6.1 Parameter Estimates

Figure 2 shows prior and posterior densities for some of the key parameters of the matrix $A(1)$ which determines the contemporaneous interaction between output, inflation, and interest rates in the unconstrained $s = 1$ regime: $\alpha_S$, $\beta_D$, $\gamma_D$, and the ratio $\gamma_D / (\alpha_S - \beta_D)$. $\alpha_S$ can be interpreted as an aggregate labor supply elasticity. Its posterior mode is approximately three. A comparison between the prior (base and adjusted) densities and the posterior
Notes: Base prior is $\tilde{p}(\theta)$ and adjusted prior is $p(\theta)$; see (26).

density indicates that the sample is not very informative about $\alpha_S$. Recall from (22) that $\alpha_S$ controls the relative response of output gap and inflation in response to a monetary policy shock. The sample is very informative about $\beta_D$ which could loosely be interpreted as the inflation coefficient in a consumption Euler equation. Recall that the (base) prior mean was 1.0 and the posterior mode is located at 0.05. The sample does not contain a lot of information about $\gamma_D$ which controls the interest rate response. According to (22) it is easier to interpret the ratio $\gamma_D/(\alpha_S - \beta_D)$ which captures the magnitude of the contemporaneous inflation movement relative to the interest rate movement in response to a monetary policy shock. The posterior mode estimate is approximately -0.1.

There are two sources of nonlinearity in our SVAR. First, the interest rate is censored. This means, even if other variables depend on the interest rate linearly, their dynamics will change nonlinearly as the interest rate reaches the ELB. Second, we allow for the laws of motion of the private-sector variables to depend on whether or not the interest rate is at or
away from the ELB. This feature of our SVAR specification capture the potential nonlinearity of agents’ decision rules in DSGE models with occasionally-binding constraints; see, for instance, ACHSV. The second nonlinearity is captured by the deviation of the parameter vector $\Phi_{12}$ introduced in Section 2.3 from zero.

Figure 3 shows the prior densities $\tilde{p}(\theta)$ and $p(\theta)$ and the posterior for the two elements of $\Phi_{12}$. The baseline prior densities $\tilde{p}(\Phi_{12,j})$ are symmetric around zero and the adjusted prior densities $p(\Phi_{12,j})$ are slightly shifted to the left. The shift is generated by the functions $f_u(\theta)$, $f_s(\theta)$, and $f_m(\theta)$. The posterior densities are clearly more concentrated than the prior densities. In terms of location, the most significant shift occurs for $\Phi_{12,2}$, which captures the kink in the inflation regression function. We will further examine the economic significance of the estimate kinks through impulse response functions.

### 6.2 Responses to Shocks

To study the effects of the two types of nonlinearities in our model and document the parameter uncertainty, we consider three types of impulse responses. First, we fix the SVAR parameters at their posterior mean values and assume that the regime $s$ persists forever. Second, we keep the parameters fixed that the posterior mean, but we allow the $s_t$ regime and hence the coefficients of the $A\Phi$ representation to switch endogenously. Because the responses now depend on the initial values of the lagged $y_t$’s we condition on two specific historical values. Third, we generate bands that reflect parameter uncertainty.

**Fixed-Regime Impulse Responses.** We begin by examining impulse responses that are computed under the assumption that the regime $s_t = s$, $s = 1, 0$, is fixed. This experiment can be interpreted as follows: the initial level $y_{1,t}^*$ is either so far above or below zero, that it does not cross zero for the next $h$ periods in a forward simulation. We also fix the parameter vector $\theta$ at its posterior mean.

Responses to minus-one-standard-deviation shocks are plotted in Figure 4. For the $s = 0$ regime, we generate impulse responses under two sets of regression functions: (i) the estimated regression functions and (ii) the $s = 1$ regression functions obtained by setting $\Phi_{12} = 0$. In the graph, the former are labeled “$s = 0$” whereas the latter are labeled “$s = 0$ (linear).” A comparison between “$s = 1$” and “$s = 0$ (linear)” highlights the effect of the interest-rate censoring at the ELB, whereas a comparison between “$s = 0$ (linear)” and $s = 0$ sheds light on the effect of the kink in the regression functions generated by the estimated $\Phi_{12}$. 
Figure 4: Responses to Negative One-Standard-Deviation Shocks with Fixed Regimes

Notes: The IRFs are computed under the assumption that the regime \( s_t = s, s = 0, 1 \), remains fixed when the desired interest rate \( y^*_t \) crosses the zero threshold. The blue lines represent the IRFs under the \( s = 1 \) regime (no censoring); the red lines are IRFs for the \( s = 0 \) regime (censoring and estimated value of \( \Phi_{12}^{\Delta} \)). The green line shows the IRFs for the \( s = 0 \) regime (censoring) under the assumption that the private-sector uses the \( s = 1 \) regression functions (i.e., \( \Phi_{12}^{\Delta} = 0 \)).

The blue lines in Figure 4 correspond to the \( s = 1 \) responses, and are computed from the implied estimates of \( \Phi(1) \) and \( \Phi'(1) \). In response to a 55 basis point (bp) expansionary monetary policy shock output gap rises by 18 bp and y-o-y inflation increases by 5 bp.
This is consistent with the posterior mode estimates of $\alpha_S$ and $\gamma_D/(\alpha_S - \beta_D)$ discussed in Section 6.1. The demand shock lowers output gap by 50 bp and inflation by 18 bp. In responses to the output and inflation drop, the central bank lowers nominal interest rates. The interest rate response is hump shaped and bottoms out at roughly -70 bp. The supply shock triggers a negative hump-shaped output-gap response with a trough at -10 bp and a 50 bp rise in inflation. The large increase in inflation leads the central bank to raise the nominal interest rate.

We now turn to the “$s = 0$ (linear)” responses (green) that solely capture the effect of censoring at the ELB. By construction, the interest rate does not react to the shocks. Nonetheless, our model generates movements in output and inflation to the monetary policy shock. For $\Phi_{12}^A = 0$, the unanticipated change in monetary policy in the ELB regime has a similar effect on output gap and inflation as in the $s = 1$ regime.

Mechanically, these responses are generated by the vector $\Phi_{12}(s_t)$ in (9). This vector is obtained for $s = 0$ from the system of equations that link the $s = 0$ regression functions to the $s = 1$ regression functions; see (16). We interpret this model feature as capturing unanticipated changes in unconventional monetary policy at the ELB. This feature of our model is shared by SVARs which use a shadow rate or a long-term interest rate (e.g., Debortoli, Galí, and Gambetti (2019)) as observable. It is also shared by specifications considered in Mavroeidis (2020) that treat the shadow rate as latent but let the private sector respond to this latent variable.

More pronounced is the effect of censoring on the responses to a demand and a supply shock. Consider a contractionary demand shock. In the $s = 1$ regime, the negative demand shock lowers output gap and inflation which generates a drop in the interest rates. In the $s = 0$ regime, interest rates cannot fall. This leads to a faster reversion of output gap and inflation to its initial level, because in general higher output gap and inflation are positively correlated with interest rates. Likewise, in response to a supply shock, the censoring amplifies the negative output gap response because it suppresses the rise in interest rates.

Finally, we turn to the $s = 0$ responses with the estimated $\Phi_{12}^A$. Even though not visible in the figure, the interest rate responses in the “$s = 0$ (linear)” regime are zero. Recall from

\[3\text{DSGE model solutions have the same feature: the monetary policy shock is a state variable and has an influence on decisions, regardless of whether or not the interest rate moves.}\]

\[4\text{Conditional on a monetary policy shocks, interest rates and output gap/inflation are negatively correlated. However, unconditionally, the monetary policy reaction function implies a positive correlation. The latter matters for Figure 4.}\]
Figure 3 that the $\Phi_{12}$ element that controls the inflation regression functions is different from zero. Hence, we expect the largest discrepancy between \(s = 0\) and \(s = 0\) (linear) to be visible in the inflation response. This is indeed the case. The monetary policy intervention is much more inflationary and the “demand” shock triggers a rise rather than a fall in inflation.\(^5\)

**Impulse Responses With Regime Shifts.** Figure 5 shows how the economy responds to minus-two-standard-deviations shocks at two different dates: 1999:Q1 and 2009:Q1. These dates are indicated by the vertical lines in Figure 1. In 1999:Q1 output gap is positive, inflation is low, and the economy is far away from the ELB. In 2009:Q1 the economy is in the midst of the Great Recession with a large negative output gap, below-mean inflation and interest rates at the ELB. As we have seen from Figure 4, the monetary policy and the demand shock tend to lower the interest rate and hence increase the probability that the ELB binds. The supply shock moves interest rates into the opposite direction, lowering the probability that the ELB binds. To generate the impulse responses, we continue to condition on posterior mean estimates of $\theta$.

The impulse responses in Figure 5 allow to the economy move toward or away from the ELB. Starting from the initial condition $x_{t+1}$, we iterate the SVAR forward to obtain two different paths of $y_{t+h}$, $h = 1, \ldots, H$. Along the baseline trajectory, we draw all innovations $\epsilon_{i,t+h}^0$ from their respective $N(0, D_{ii})$ distributions. We denote the resulting series $y_{t+1,t+H}^0$. To generate the shocked path $y_{t+1,t+H}^1$, we set $\epsilon_{1,t+1}^1 = \epsilon_{1,t+1}^0 + \delta$ and let $\epsilon_{i,t+h}^1 = \epsilon_{i,t+1}^0$ for all other $(i, h)$. Here $\delta$ is the size of the shock. The impulse response is defined as the difference between the shocked and the baseline path. The simulation is repeated for $j = 1, \ldots, M$ where we use $M = 10,000$.

In the top three rows of Figure 5 we plot the distribution of $y_{t+1,t+H}^j - y_{t+1,t+H}^0$ across $j$. The bands depict 90% equal-tail-probability intervals and the solid lines are means. In the absence of the nonlinearities triggered by the interest rate reaching the ELB, the bands collapse to a single line. The bottom panel of Figure 5 shows bands for the interest rate path along the “shocked” trajectories. Each time, a trajectory $j$ hits the ELB, the difference between the shocked path and the baseline path will start to deviate from the fixed-regime $s = 1$ impulse response. Once a sufficiently large number of trajectories have hit the ELB, the response bands will start to fan out. Because of this cumulative effect, it is not necessary that the fanning out of the response bands coincides with the 5% quantile of the interest rate.

\(^5\)The sign reversal of the inflation response relative to the $s = 1$ regime is also visible in the DSGE model based impulse responses reported in Figure 5 of Aruoba, Cuba-Borda, and Schorfheide (2018). Beyond that, the response functions are not directly comparable because the shock identification is different.
Figure 5: Responses in 1999:Q1 and 2009:Q1: Flexible Regimes and Shock Uncertainty

Notes: The top 3 panels show the distribution of the difference between a simulated baseline trajectory and a trajectory in which a two-standard-deviation shock is subtracted from the simulated monetary policy / demand / supply shock in the initial period. The bottom panels show the level of the interest rate under the “shocked” trajectory. Bands indicate 5% and 95% quantiles and solid lines are means across $M = 10,000$ distribution represented by the lower bound of the grey bands in the last row of Figure 5 reaching zero.
Qualitatively, the 1999:Q1 responses look very similar to the $s = 1$ responses in Figure 5 because on impact the nominal interest is approximately 5% and the economy is far way from the ELB. However, the size of the shocks and thus the responses are twice as large. For the monetary policy shock the, impulse response bands collapse to a single line because it is very unlikely that the economy reaches the ELB along the “shocked” and the baseline paths.

For the demand and supply shocks the mean responses in 1999:Q1 look similar to the $s = 1$ responses in Figure 4, except that the former are twice as large. For the demand shock responses the band starts to widen after horizon $h = 10$. This means that there are a substantial number of trajectories along which interest rates are strictly positive under the baseline scenario ($s = 1$), but hit the ELB under the shocked scenario ($s = 0$). Along these trajectories, the interest rate response reverts more quickly back to zero (because the interest rate drop is constrained by the ELB), and so does output gap and the inflation response. Visually, in Figure 5 the response along the upper end of the blue bands starts to look more similar to the mean response under the 2009:Q1 initial conditions.

Conditional on the 2009:Q1 initial conditions, the responses broadly resemble the $s = 0$ responses in Figure 4 because the U.S. interest rates have reached the ELB. A closer look reveals that the IRFs are slightly tilted toward the $s = 1$ responses because under some of the simulated trajectories, the economy quickly moves away from the ELB.

**Parameter Uncertainty.** We will now explore the posterior uncertainty associated with the impulse response functions.\(^6\) The bands reported in Figure 6 represent 90% equal-tail-probability credible intervals that reflect posterior parameter uncertainty. The solid lines are pointwise medians of the impulse response posteriors. As in the previous figure, we compare responses based on the 1999:Q1 (blue) and 2009:Q1 (red) initial conditions. Figure 6 is generated by converting draws $\theta^i$ from the posterior distribution into mean responses (each computed over 10,000 simulations) to the three structural shocks depicted as solid lines in Figure 5.

A comparison of the 1999:Q1 and 2009:Q1 bands sheds light on the question whether the propagation of shocks is different at the ELB. In regard to interest rate responses, the

---

\(^6\)A comparison of prior and posterior impulse response bands is provided in the Online Appendix. The prior bands are substantially wider than the posterior bands indicating that the sample is informative about the propagation of shocks. The Online Appendix also provides a comparison of responses from a linear version of our SVAR to the responses reported in BH. Our responses to supply and demand shocks are both qualitatively and quantitatively similar, except that for our model specification there is less posterior uncertainty about the impact effect of the shocks. The main difference between the monetary policy responses is that in our case both output gap and inflation react less strongly than in the BH estimates.
answer is a trivial yes, because it is directly constrained by the ELB. More interesting is the comparison for output gap and inflation. Consistent with the insignificant estimate of $\Phi_{12}^\Delta$ for the output gap equation reported in Figure 3, we see that the output gap impulse response bands strongly overlap for the two sets of initial conditions.

The inflation responses to a monetary policy and a demand shock, on the other hands, show a substantial difference with little overlap of the two bands in the periods right after the shock. The unconventional expansionary monetary policy intervention at the ELB in 2009:Q1 is much more inflationary than away from the ELB in 1999:Q1. Moreover, the price adjustment in response to a demand shock is much more muted in 2009:Q1 and has the opposite sign than in 1999:Q1. Thus, we conclude that, as predicted by DSGE models with an ELB constraint, the ELB is not irrelevant for the propagation of shocks and this effect is measurable in a parsimonious SVAR framework that allows for changes in the private-sector
behavior at the ELB.

6.3 Shadow Rate

Our model allows us to generate a shadow rate. It is given by an estimate of the latent variable $y_{1,t}^*$. In the Online Appendix we provide a formula for $p(y_{1,t}^*|y_{1,t} = 0, y_{2,t}, x_t, \theta)$. Because in our SVAR $x_t$ depends on $y_{1,t-1:t-p}$ instead of $y_{1,t-1:t-p}^*$ inference on the shadow rate is static and does not require dynamic filtering and smoothing. The posterior median estimate, 60% and 90% bands of the model-implied distribution of the shadow rate are plotted in Figure 7. The bands reflect uncertainty about the parameters $\theta$ and the shocks conditional on $(y_{1,t} = 0, y_{2,t}, x_t, \theta)$. The inference exploits the correlation structure implied by $\Phi(0)$ and $\Sigma(0)$ between interest rates, on the one hand, and output gap and inflation on the other hand.

The estimated shadow rate drops to -3% in 2009 during the first large scale asset purchase intervention (QE1) of the Federal Reserve. From 2010 to 2015, during QE2, Operation Twist, and QE3 it hovered around -1%. This shows that the Federal Reserve was particularly aggressive in providing monetary stimulus in 2009 and reverted to a lesser but consistent stimulus for the rest of the ELB episode. It is also noteworthy that the slope of the shadow rate in 2009 until the official end of the recession in 2009Q2 closely matches the slope in the federal funds rate just prior to the start of the ELB episode. Qualitatively, the time path of the shadow rate is consistent with the time path of the desired interest rate (red solid line) estimated with a small-scale New Keynesian DSGE in Aruoba, Cuba-Borda, and Schorfheide (2018). Quantitatively, the trough in the DSGE model implied shadow rate occurred about six months after the trough in the SVAR based shadow rate.

Finally, we provide a comparison with the yield-curve based shadow rate of Wu and Xia (2016). The Wu-Xia rate, rather than being based on a censored interest-rate feedback rule, is based on a censored affine term structure model and extracts information from yields on medium- and long-term bonds. Their shadow rate troughs in 2014, about five years after the end of the Great Recession, about a year prior to the lift-off from the ELB. We find this somewhat implausible in view of the narrative evidence that the most significant interventions happened during and right after the Great Recession - which is consistent with our SVAR based estimates.
Notes: Blue line outside of the ELB episode: federal funds rate. During the ELB episode: 60% (green shade) and 90% (blue shade) equal-tail-probability bands. Black line is Wu-Xia shadow rate and red line is the desired interest rate from the DSGE model of Aruoba, Cuba-Borda, and Schorfheide (2018). The solid vertical line denotes September 2008. The yellow shading indicates the months when the particular Fed program was active. During the green shaded area, both Operation Twist and QE3 were active. The dashed-dotted vertical line shows January 2012, when the formal inflation target was announced. The dashed vertical line shows the “taper tantrum” episode in June 2013.

7 Conclusion

We developed a structural VAR in which an occasionally-binding constraint generates censoring of one of the dependent variables. Once the censoring mechanism is triggered, we allow some of the coefficients for the remaining variables to change. By imposing that the regression functions are continuous at the censoring point, we can show that under some mild parameter restrictions delivers a unique reduced form. The resulting model is more parsimonious than a time-varying-coefficient VAR and the switch in parameter values is tied to the censoring mechanism, which in our application is the ELB on nominal interest rates. An application to U.S. data provided evidence of a significant shift of parameters in the inflation equation. In future work we plan to estimate the model based on data from other countries that reached the ELB.
References


Online Appendix to “SVARs With Occasionally-Binding Constraints”

S. Borağan Aruoba, Frank Schorfheide, and Sergio Villalvazo

This Appendix consists of the following sections:

A. Proofs and Theoretical Derivations
B. Data Sources
C. Additional Empirical Results
A Proofs and Theoretical Derivations

A.1 Proof of Proposition 1

We will drop the time subscripts. Because of the linearity of the $s=0$ and $s=1$ regression functions, the functions $x'\Phi_1(s=0) + u_1(s=0,\epsilon)$ and $x'\Phi_1(s=1) + u_1(s=1,\epsilon)$ are continuous in $(x,\epsilon)$. We regard $x$ as predetermined. Given $x$, let $\mathcal{E}^0(x)$ denote the set of innovations for which the ELB becomes binding. For any $\epsilon_0 \in \mathcal{E}^0(x)$, the following restriction holds under continuity:

$$x'\Phi_1(0) + u_1(0,\epsilon_0) = 0 \quad \text{and} \quad x'\Phi_1(1) + u_1(1,\epsilon_0) = 0.$$  

The proposition is proved if we can show that for any $\epsilon \not\in \mathcal{E}^0(x)$ the following conditions hold:

$$x'\Phi_1(1) + u_1(1,\epsilon) > 0 \quad \text{implies} \quad x'\Phi_1(0) + u_1(0,\epsilon) > 0$$

$$x'\Phi_1(0) + u_1(0,\epsilon) \leq 0 \quad \text{implies} \quad x'\Phi_1(1) + u_1(1,\epsilon) \leq 0.$$  

We will assume that $A_{11} > 0$ (otherwise inequalities will be reversed).

Define the slackness

$$\Delta(s,\epsilon) = x'\Phi_1(s) + u_1(s,\epsilon)$$

$$= \left[ x_1(B_{11} - \Phi_{12}(s)A_{21}) + x_2'(B_{21} - \Phi_{22}(s)A_{21}) \right] + \left[ \epsilon_1(1 - \Phi_{12}'(s)A_{21}) - \epsilon_2'\Phi_{22}'(s)A_{21} \right]$$

$$= \Delta_x(s,\epsilon) + \Delta_e(s,\epsilon).$$

Using this notation, we need to show that

$$\text{sign}(\Delta(1,\epsilon)) = \text{sign}(\Delta(0,\epsilon)). \quad (A.1)$$

We will derive a functional relationship between $\Delta$ terms for $s=1$ and $s=0$. Using the
continuity restrictions derived above, we can write

\[
\Delta_x(0, \epsilon) = x_1(B_{11} - \Phi_{12}(0)A_{21}) + x'_2(B_{21} - \Phi_{22}(0)A_{21})
\]
\[
= x_1(B_{11} - \Phi_{12}(0)A_{21}) + x'_2\left(B_{21} - \left[\Phi_{22}(1) + \frac{B_{21} - \Phi_{22}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} \Phi_{12}\right]A_{21}\right)
\]
\[
= x_1(B_{11} - \Phi_{12}(0)A_{21}) + x'_2(B_{21} - \Phi_{22}(1)A_{21}) - x'_2\frac{B_{21} - \Phi_{22}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} \Phi_{12}A_{21}
\]
\[
= x_1(B_{11} - (\Phi_{12}(1) + \Phi_{12}A_{21}) + x'_2(B_{21} - \Phi_{22}(1)A_{21}) - x'_2\frac{B_{21} - \Phi_{22}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} \Phi_{12}A_{21}
\]
\[
= \Delta_x(1, \epsilon) - x_1 \Phi_{12}A_{21} + x'_2\frac{B_{21} - \Phi_{22}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} \Phi_{12}A_{21}
\]
\[
= \Delta_x(1, \epsilon) - x_1 \frac{B_{11} - \Phi_{12}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} \Phi_{12}A_{21} - x'_2\frac{B_{21} - \Phi_{22}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} \Phi_{12}A_{21}
\]
\[
= \Delta_x(1, \epsilon) - \left(x_1(B_{11} - \Phi_{12}(1)A_{21}) + x'_2(B_{21} - \Phi_{22}(1)A_{21})\right)\frac{\Phi_{12}A_{21}}{B_{11} - \Phi_{12}(1)A_{21}}
\]
\[
= \Delta_x(1, \epsilon) \left(1 - \frac{\Phi_{12}A_{21}}{B_{11} - \Phi_{12}(1)A_{21}}\right).
\]

Moreover,

\[
\Delta_\epsilon(0, \epsilon) = \epsilon_1(1 - \Phi_{12}(0)A_{21}) - \epsilon'_2\Phi_{22}(0)A_{21}
\]
\[
= \epsilon_1\left(1 - \left(\Phi_{12}(1) + \frac{1 - \Phi_{12}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} \Phi_{12}\right)A_{21}\right)
\]
\[-\epsilon'_2\left(\Phi_{22}(1) - \frac{\Phi_{22}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} \Phi_{12}\right)A_{21}
\]
\[
= \epsilon_1\left(1 - \Phi_{12}(1)A_{21}\right) - \epsilon'_2\Phi_{22}(1)A_{21}
\]
\[-\epsilon_1\frac{1 - \Phi_{12}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} \Phi_{12}A_{21} + \epsilon'_2\frac{\Phi_{22}(1)A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} \Phi_{12}A_{21}
\]
\[
= \Delta_\epsilon(1, \epsilon) - \left(\epsilon_1(1 - \Phi_{12}(1)A_{21}) - \epsilon'_2\Phi_{22}(1)A_{21}\right)\frac{\Phi_{12}A_{21}}{B_{11} - \Phi_{12}(1)A_{21}}
\]
\[
= \Delta_\epsilon(1, \epsilon) \left(1 - \frac{\Phi_{12}A_{21}}{B_{11} - \Phi_{12}(1)A_{21}}\right).
\]

This leads to

\[
\Delta(0, \epsilon) = \Delta(1, \epsilon) \left(1 - \frac{\Phi_{12}A_{21}}{B_{11} - \Phi_{12}(1)A_{21}}\right).
\]
We deduce that (A.1) holds iff
\[ 1 - \frac{\Phi_{12}A_{21}}{B_{11} - \Phi_{12}(1)A_{21}} > 0. \] (A.2)

Note that this condition neither depends on \( x \) nor on \( \epsilon \). Here \( B_{11} \) is the reaction of interest rates to the lagged interest rate \( x_{1,t} \), \( \Phi_{12}(0) \) is the reaction of the private-sector variables to the lagged interest rate \( x_{1,t} \), and \( A_{21} \) is the response of the interest rate to the contemporaneous private-sector variables \( (y_{2,t}) \).

Recall from (13) that the ELB starts to bind when
\[ x_{1,t}(B_{11} - \Phi_{12}(1)A_{21}) = -x'_{2,t}(B_{21} - \Phi_{22}(1)A_{21}) - \epsilon_{1,t}(1 - \Phi'_{12}(1)A_{21}) + \epsilon'_{2,t}(\Phi'_{22}(1)A_{21}). \]

Denote the coefficient on \( x_{1,t} \) by \( D = B_{11} - \Phi_{12}(1)A_{21} \). We distinguish two cases. First, suppose that \( D > 0 \). Then multiplying (A.2) by \( D \) leads to the inequality
\[ D - \Phi_{12}A_{21} > 0 \quad \text{or} \quad D > \Phi_{12}A_{21}. \]

Second, suppose that \( D < 0 \). Then multiplying (A.2) by \( D \) leads to the inequality
\[ D - \Phi_{12}A_{21} < 0 \quad \text{or} \quad D < \Phi_{12}A_{21}. \]

By writing
\[ D - \Phi_{12}A_{21} = B_{11} - \Phi_{12}(1)A_{21} - (\Phi_{12}(0) - \Phi_{12}(1))A_{21} = B_{11} - \Phi_{12}(0)A_{21} \]
we obtain the expressions in the statement of the proposition. □

A.2 Derivations for the Likelihood Function

The derivations for \( p(y_{1,t}|\cdot) \) in (31) and \( p(y_{2,t}|y_{1,t} > 0, \cdot) \) in (33) are straightforward. We will focus on the derivation of \( p(y_{2,t}|y_{1,t} = 0, \cdot) \) in (32). In the latter case, the specific value of \( u_{1,t}(0) \) is unknown. Write
\[ p(y_{2,t}|y_{1,t} = 0, \cdot) = \int p(y_{2,t}|u_{1,t})p(u_{1,t}|y_{1,t} = 0, \cdot)du_{1,t}. \] (A.3)
We begin by deriving the density of $u_{1,t}$ conditional on $y_{1,t} = 0$:

$$
p(u_{1,t}|y_{1,t} = 0, \cdot) = \frac{p_N(u_{1,t}; 0, \Sigma_{11}(0)) \mathbb{I}\{u_{1,t}(0) \leq -x'_t\Phi_1(0)\}}{F_N\left(-\frac{x'_t\Phi_1(0)}{\sqrt{\Sigma_{11}(0)}}\right)}. \quad (A.4)
$$

Combining (A.3) and (A.4) yields:

$$
p(y_{2,t}|y_{1,t} = 0, \cdot) = \left[F_N\left(-\frac{x'_t\Phi_1(1)}{\sqrt{\Sigma_{11}(1)}}\right)\right]^{-1} \int p_N(u_{2,t}(0); u_{1,t}M_{12}(0), \Sigma_{21}(0)) \times p_N(u_{1,t}; 0, \Sigma_{11}(0)) \mathbb{I}\{u_{1,t}(0) \leq -x'_t\Phi_1(0)\} du_{1,t}. \quad (A.5)
$$

In the subsequent steps we will evaluate the integral.

To simplify the notation we drop the $s_t = 0$ argument from $\Sigma$ and $\Phi$ matrices. The key terms that appear in the integral in (A.5) can be manipulated as follows:

$$
\begin{align*}
(u'_t - u_{1,t}M_{12})\Sigma_{21}^{-1}(u_{2,t} - M_{21}u_{1,t}) + u_{1,t}\Sigma^{-1}_{11}u_{1,t} \\
= u'_t\Sigma_{21}^{-1}u_{2,t} - 2u_{1,t}M_{12}\Sigma_{21}^{-1}u_{2,t} + u_{1,t}M_{12}\Sigma_{21}^{-1}M_{21}u_{1,t} + u_{1,t}\Sigma_{11}^{-1}u_{1,t} \\
= u'_t\Sigma_{21}^{-1}u_{2,t} + (M_{12}\Sigma_{21}^{-1}M_{21} + \Sigma_{11}^{-1})\left(u_{1,t} - \frac{M_{12}\Sigma_{21}^{-1}u_{2,t}}{M_{12}\Sigma_{21}^{-1}M_{21} + \Sigma_{11}^{-1}}\right)^2 \\
- \frac{(M_{12}\Sigma_{21}^{-1}u_{2,t})^2}{M_{12}\Sigma_{21}^{-1}M_{21} + \Sigma_{11}^{-1}}.
\end{align*}
$$

Define

$$
\tilde{M}_u = \frac{M_{12}\Sigma_{21}^{-1}}{M_{12}\Sigma_{21}^{-1}M_{21} + \Sigma_{11}^{-1}}, \quad \tilde{V}_u = (M_{12}\Sigma_{21}^{-1}M_{21} + \Sigma_{11}^{-1})^{-1}
$$

Notice that

$$
(2\pi)^{-1/2}|\tilde{V}_u|^{-1/2} \times \int \exp\left\{-\frac{1}{2\tilde{V}_u}(u_{1,t} - \tilde{M}_u u_{2,t})^2\right\} \mathbb{I}\left\{\frac{u_{1,t} - \tilde{M}_u u_{2,t}}{\sqrt{\tilde{V}_u}} \leq -\frac{x'_t\Phi_1 + \tilde{M}_u u_{2,t}}{\sqrt{\tilde{V}_u}}\right\} du_{1,t}
$$

$$
= F_N\left(-\frac{x'_t\Phi_1 + \tilde{M}_u u_{2,t}}{\sqrt{\tilde{V}_u}}\right).
$$
Therefore, we obtain Equation (33) in the main text:

\[ p(y_{2,t} | y_{1,t} = 0, \cdot) = (2\pi)^{-(n-1)/2} |\Sigma_{21}(0)|^{-1/2} |\Sigma_{11}(0)|^{-1/2} |\tilde{V}_u(0)|^{1/2} \left[ F_N \left( -\frac{x_t' \Phi_1(1)}{\sqrt{\Sigma_{11}(1)}} \right) \right]^{-1} \times F_N \left( -\frac{x_t' \Phi_1(0) + \tilde{M}_u(0) u_{2,t}(0)}{\sqrt{\tilde{V}_u(0)}} \right) \times \exp \left\{ -\frac{1}{2} u_{2,t}' \Sigma_{21}^{-1}(0) u_{2,t}(0) + \frac{1}{2} \tilde{V}_u^{-1}(0) [\tilde{M}_u(0) u_{2,t}(0)]^2 \right\} \]

\[ \square \]

A.3 Derivations for the Shadow Rate

The shadow rate can be characterized through the density \( p(y_{1,t}^* | y_{1,t} = 0, y_{2,t}, x_t, \theta) \). Recall that our model conditional on \( y_{1,t} = 0 \) (and, hence, \( s_t = 0 \)) has the reduced-form representation

\[ y_{1,t}^* = x_t' \Phi_1(0) + u_{1,t}(0), \quad y_{2,t}^* = x_t' \Phi_2(0) + u_{2,t}'(0). \]  

(A.6)

Define

\[ \Sigma_{12}(0) = \Sigma_{11}(0) - \Sigma_{12}(0) \Sigma_{22}^{-1}(0) \Sigma_{21}(0), \quad M_{21}(0) = \Sigma_{22}^{-1}(0) \Sigma_{21}(0). \]  

(A.7)

Thus,

\[ u_{1,t} | u_{2,t} \sim N(u_{2,t} M_{21}(0), \Sigma_{12}(0)). \]

Now condition on \( y_{1,t} = 0 \) which we can express as \( u_{1,t}(0) \leq -x_t' \Phi_1(0) \):

\[ p(u_{1,t} | y_{1,t} = 0, u_{2,t}(0), \cdot) = \frac{p_N(u_{1,t}; u_{2,t} M_{21}(0), \Sigma_{12}(0)) \mathbb{I}\{u_{1,t}(0) \leq -x_t' \Phi_1(0)\}}{F_N \left( -\frac{x_t' \Phi_1(0) - u_{2,t}' M_{21}(0)}{\sqrt{\Sigma_{12}(0)}} \right)}. \]  

(A.8)

To generate a draw from the posterior of the shadow rate, we can proceed as follows:

(i) Compute \( u_{2,t}(0) = y_{2,t}^* - x_t' \Phi_2(0) \).

(ii) Compute \( \Sigma_{12}(0) \) and \( M_{21}(0) \) using (A.7).

(iii) Draw \( \tilde{u}_{1,t} \sim p(u_{1,t} | y_{1,t} = 0, u_{2,t}(0), \cdot) \).

(iv) Compute \( \tilde{y}_{1,t}^* = x_t' \Phi_1(0) + \tilde{u}_{1,t} \).

Because \( x_t \) depends on \( y_{1,t-1:t-p} \) instead of \( y_{1,t-1:t-p}^* \) inference on the shadow rate is static and does not require dynamic filtering.
B Data Sources

All series are obtained from the economic research database of the Federal Reserve Bank of St. Louis (FRED). Real GDP is \( gdpc1 \), real potential GDP is \( gdppot \), the personal consumption expenditure deflator is \( dpcerd3q086sbea \), and the federal funds rate is \( fedfunds \). We average the monthly rates to time-aggregate the interest rates to quarterly frequency.

C Additional Empirical Results

C.1 Further Results from the ELB-SVAR

In Figures A-1 and A-2 we compare prior (grey) and posterior (green) uncertainty about the impulse responses generated conditional on the states of the economy in 1999:Q1 and 2009:Q1, respectively. Our prior distribution allows for a wide variety of responses to the three shocks. While the prior distribution of the impact effect is constrained by the sign restrictions imposed on \( \alpha_S \), \( \beta_D \), and \( \gamma_D \), the wide bands for \( h \geq 2 \) indicate that the prior leaves the sign of the dynamic effect largely unconstrained. The posterior distribution of the impulse responses is substantially more concentrated, reflecting the sample information about the effects of the three shocks.
Figure A-1: Responses to Minus-Two-Standard-Deviation Shocks with Flexible Regimes and Prior/Posterior Parameter Uncertainty in 1999:Q1

Notes: Prior median (black) and 90% equal-tail-probability credible interval (grey); posterior median (green) and 90% credible interval (light green).
Figure A-2: Responses to Minus-Two-Standard-Deviation Shocks with Flexible Regimes and Prior/Posterior Parameter Uncertainty in 2009:Q1

Notes: Prior median (black) and 90% equal-tail-probability credible interval (grey); posterior median (green) and 90% credible interval (light green).
C.2 Results From a Linear SVAR

With the exception of the censoring and the kinked regression functions to capture the ELB episode in the U.S. our SVAR specification is very similar to the one estimated by Baumeister and Hamilton (2018), henceforth BH. For comparison purposes, we estimate a version of the SVAR without censoring and kinks in the regression functions using the BH sample, which ranges from 1986:Q1 to 2008:Q3. Our parameterization of the $A$ matrix is identical to BH, but we are using a slightly different prior distribution for its elements, in particular for $\alpha_S$, $\beta_D$, and $\gamma_D$. Because the SVAR is only set-identified, the posterior distribution is in some dimensions quite sensitive to the choice of the prior. Moreover, we use a VAR with one lag, whereas BH use a VAR with 4 lags.

We convert draws from the prior and the posterior distribution into impulse response functions and plot posterior means as well as 90% equal-tail-probability bands in Figure A-3 which can be compared to Figure 4 of BH. Our responses to supply and demand shocks are both qualitatively and quantitatively similar, except that for our model specification there is less posterior uncertainty about the impact effect of the shocks. The main difference between the monetary policy responses is that in our case both output gap and inflation react less strongly than in the BH estimates.
Figure A-3: Response to One-Standard-Deviation Shocks, Linear SVAR

IRFs from (Adjusted) Prior Draws

IRFs from Posterior Draws

Notes: The solid lines indicate the means of the impulse response functions and the bands represent 90% equal-tail-probability credible intervals.