SHOCK RESTRICTED STRUCTURAL VECTOR-AUTOREGRESSIONS

SYDNEY C. LUDVIGSON[*]

NYU and NBER

SAI MA[†]

Federal Reserve Board

SERENA NG [‡]

Columbia University and NBER

December 11, 2019

Abstract

It is well known that the covariance structure of the data alone is not enough to identify an SVAR, and the conventional approach is to impose restrictions on the parameters of the model based on a priori theoretical considerations. This paper suggests that much can be gained by requiring the properties of the identified shocks to agree with major economic events that have been realized. We first show that even without additional restrictions, the data alone are often quite informative about the quantitatively important shocks that have occurred in the sample. We propose shrinking the set of solutions by imposing two types of inequality constraints on the shocks. The first restricts the sign and possibly magnitude of the shocks during unusual episodes in history. The second restricts the correlation between the shocks and variables external to the SVAR. The methodology provides a way to assess the validity of assumptions imposed as equality constraints. The effectiveness and limitations of this approach are exemplified with three applications.

JEL: C13, C18, C26, C36

Keywords: Plausibly exogenous regressors, large shocks, external instruments, VAR.

1 Introduction

A challenge in economic analysis is that the data represented by a vector autoregression (VAR) can be consistent with many causal structures differentiated by distinct economic models and primitive assumptions. Structural vector-autoregressive models (SVARs) provide a simple framework that enables researchers to perform counter-factual experiments without fully characterizing all primitives or micro-foundations that lead to the dynamic system. Concisely stated, a n-variable SVAR analysis consists of finding an $n \times n$ matrix $\mathbf{B}$ that relates the reduced-form innovations $\boldsymbol{\eta}_t$ to the mutually uncorrelated structural shocks $\mathbf{e}_t$:

$$\boldsymbol{\eta}_t = \mathbf{B}\mathbf{e}_t.$$

The data provide $\boldsymbol{\eta}_t$, but the above relationship only gives $n(n+1)/2$ pieces of information about $\mathbf{B}$ through the reduced-form covariance restrictions. Hence $n(n-1)/2$ additional restrictions are necessary to identify $\mathbf{B}$. Sims (1980) originally proposed a triangularized system to identify $\mathbf{B}$, but long-run economic restrictions, statistical restrictions based on heteroskedasticity of the VAR innovations, or additional information in the form of high frequency data, variables external to the SVAR, or narrative descriptions have also been employed.[1] Point identification exists if these restrictions are enough to yield a unique solution. In some cases, point identification is achievable only by imposing restrictions that are difficult to defend. It may then be desirable to abandon the goal of point identification in favor of less restrictive economic assumptions.

In this paper we are concerned with applications for which economic restrictions that permit point identification are not available. Though the literature has focused on finding $n(n-1)/2$ restrictions by appealing to theory or prior information, what seems to have been overlooked is the information conveyed by the $n(n+1)/2$ atheoretical restrictions. Though they only provide an under-identified set of solutions, the distribution of shocks in this set is telling about the events that might have happened, and is a useful guide to the additional restrictions that need to be imposed. Unlike a typical SVAR which puts restrictions on the dynamic responses or propagating mechanism, we put restrictions on the shocks. For this reason, we refer to our approach as a "shock-restricted" SVARs.

Although many types of shock-based restrictions are possible, we focus here on two main types. The first is a set of inequality restrictions that require the identified shocks to have defensible properties during special episodes of history for which a broad historical understanding would suggest a certain behavior of the structural shocks. Historical narratives are augmented and verified by using the data to locate distinguishing characteristics of the structural shocks. We refer to these as *event inequality constraints*. The second are a set of inequality restrictions that require the identified shocks to exhibit a non-zero correlation with certain variables exter-

---

[1] For a comprehensive review of SVAR models, see Ramey (2016), Kilian and Lutkepohl (2016).

nal to the VAR that should be informative about the shocks of interest. We refer to these as *external variable inequality constraints.* Many restrictions used in econometric modeling can be written as external variable equality constraints. We weaken these assumptions by allowing for inequalities.

The use of inequality constraints has a long history in the SVAR literature. They have predominantly been formulated as sign restrictions on impulse responses with zero as the threshold value, e.g., Uhlig (2005). We entertain inequality constraints on the shocks, either on their behavior during particular episodes or in terms of their comovement with external variables, possibly with non-zero thresholds. Large threshold values allow us to isolate and exploit the information in rare events. Working with inequalities comes at the cost of foregoing point identification, but the approach nonetheless enables us to check whether shocks thought to have happened can be recovered under the assumptions of the model. The approach also allows us to evaluate whether assumptions such as exogeneity, zero, and unit elasticity, or others are valid.

The methodology described here is predicated on the idea that a credible identification scheme should produce estimates of $\mathbf{e}_t$ that are congruent with our ex-post understanding of historical events and/or with broadly accepted economic notions of a shock's defining properties. For example, a scheme that identifies a large positive output shock in the 1982 recession would be dismissed because the existence of a shock would be hard to defend given the historical account of the events at the time. Similarly, a scheme that produces uncertainty shocks that are negatively rather than positively correlated with the value of safe haven assets would be dismissed given an economic understanding of uncertainty as a phenomenon that enhances rather than weakens precautionary motives. Such shock-based restrictions turn out to be valuable for identification because, although two feasible structural models $\mathbf{B}$ and $\tilde{\mathbf{B}}$ will generate shocks $\mathbf{e}_t$ and $\tilde{\mathbf{e}}_t$ with equivalent first and second moments, the $\mathbf{e}_t$ and $\tilde{\mathbf{e}}_t$ are not necessarily the same at any given t or in terms of their comovement with external proxy variables. In other words, two series with equivalent properties "on average" can still have distinguishable features in certain subsamples and/or in their comovement with external variables.

At a superficial level, the shock-based constraints appear similar to identification schemes already present in the literature. Many important studies have used a narrative approach to construct shock series from historical readings of political and economic events. These shock series are typically used in an SVAR context as an external instrumental variable. A recent literature pioneered by Mertens and Ravn (2013) and Stock and Watson (2008) has emerged that uses economic time series external to the SVAR to help with identification. These approaches achieve point identification by assuming that the external variables have a zero correlation with some shocks (an exogeneity assumption) and is strongly correlated with other shocks (a relevance assumption). By contrast, as in Conley, Hansen and Rossi (2012), our methodology allows the external variables and events to exhibit departures from exogeneity. But while their

approach is Bayesian and focuses on single equation estimation, our approach is frequentist in the spirit of the moment inequality framework of Andrews and Soares (2010) and focuses on system estimation. And unlike Conley et al. (2012), the motivation for the event and external variable inequality constraints considered here is not limited to relaxing exogeneity assumptions; the more general objective is to use them to help with identification.

Like any identification scheme, the one studied here is not without limitations. While the approach often allows for weaker assumptions, clear conclusions may not emerge without the imposition of multiple shock-based restrictions, each of which need to be located and defended. This in turn requires a detailed reading of the relevant events. Moreover, alternative combinations of these constraints, as well as alternative parameterizations of a given set of constraints, may produce differing results. In such cases, a sensitivity analysis may be undertaken to reveal how fragile the findings are to different identifying assumptions.

In what follows, we demonstrate how the methodology can be used with three applications. In the first, building off of previous work by Kilian (2008) and Kilian and Murphy (2012), we use an SVAR for the oil market to illustrate how to set up the restrictions and obtain a set of plausible solutions. It is shown that, even though the dynamic responses under different restrictions are similar, some shocks thought to have occurred in particular episodes are not actually evident under restrictions previously used in the literature. In the second application, we consider the SVAR used in Gertler and Karadi (2015) to assess the transmission of monetary policy shocks. Results using our shock-based restrictions broadly support their assumption about instrument exogeneity, but reveal that periods of monetary easing are the source of the identifying power. Finally, inspired by the work of Mian and Sufi (2014), we use a bivariate SVAR to estimate the macroeconomic effects of fluctuations in housing wealth on aggregate consumption. We find that much can be said about these effects even in the absence of a credibly valid instrument, and that the shock-based restrictions produce tighter bounds than competing methods in the literature. A fourth example, given in Ludvigson, Ma and Ng (2019) and hereafter referred to as LMN, is used to distinguish first from second moment shocks, a problem for which theory offers little guidance. These different applications illustrate how the event and external variable inequality constraints may be used together or separately, and possibly in conjunction with other identification schemes that already exist in the literature.

A general finding is that there is much to be gained by requiring the properties of the identified shocks to agree with major economic events that have been realized. Often, the shock-based restrictions are rich enough to conclude that the data are consistent with a clear causal pattern among the variables. By contrast, the common approach of focusing exclusively on restricting the SVAR parameters to achieve identification often misses valuable information about the structural shocks of interest. In some cases, the shocks implied by identification schemes derived under conventional parameter restrictions appear implausible given an ex post

understanding of major historical or economic events.

The rest of this paper is organized as follows. The next section describes the econometric framework and the shock-based restrictions at a general level. Sections 3, 4 and 5 present an analysis of the three applications mentioned above. Section 6 concludes.

2 Econometric Framework

Let $\mathbf{X}_t$ denote a $n \times 1$ vector time series. We suppose that $\mathbf{X}_t$ has a reduced-form finite-order autoregressive representation $\mathbf{X}_t = \sum_{j=1}^{p} \mathbf{A}_j \mathbf{X}_{t-j} + \boldsymbol{\eta}_t$, $\boldsymbol{\eta}_t \sim (0, \boldsymbol{\Omega})$, $\quad \boldsymbol{\Omega} = \mathbf{P}\mathbf{P}'$ where $\mathbf{P}$ is the unique lower-triangular Cholesky factor with non-negative diagonal elements. The reduced-form parameters are collected into $\boldsymbol{\phi} = \left(\text{vec}(\mathbf{A}_1)' \quad \ldots \quad \text{vec}(\mathbf{A}_p)', \quad \text{vech}(\boldsymbol{\Omega})' \right)'$. The reduced-form innovations $\boldsymbol{\eta}_t$ are related to the structural SVAR shocks $\mathbf{e}_t$ by an invertible matrix $\mathbf{H}$:

$$\boldsymbol{\eta}_t = \mathbf{H}\boldsymbol{\Sigma}\mathbf{e}_t \equiv \mathbf{B}\mathbf{e}_t, \qquad \mathbf{e}_t \sim (0, \mathbf{I_K}), \qquad \text{diag}\,(\mathbf{H}) = 1,$$

where $\mathbf{B} \equiv \mathbf{H}\boldsymbol{\Sigma}$, and $\boldsymbol{\Sigma}$ is a diagonal matrix with variance of the shocks in the diagonal entries. The structural shocks $\mathbf{e}_t$ are mean zero with unit variance, serially and mutually uncorrelated. We adopt the unit effect normalization that $H_{jj} = 1$ for all j.

The goal of the exercise is analyze the dynamic effects of $\mathbf{e}_t$ on $\mathbf{X}_t$. Let "hats" denote estimated variables. Since the autoregressive parameters $\mathbf{A}_j$ can be consistently estimated under regularity conditions, the sample residuals $\hat{\boldsymbol{\eta}}_t(\hat{\boldsymbol{\phi}})$ are consistent estimates of $\boldsymbol{\eta}_t$. The empirical SVAR problem reduces to finding $\mathbf{B}$ from $\hat{\boldsymbol{\phi}}$. But there are n^2 parameters in $\mathbf{B}$ and the sample covariance of $\hat{\boldsymbol{\eta}}_t$ only provides $n(n+1)/2 < n^2$ conditions $\bar{g}_Z(\mathbf{B})$ in the form

$$\bar{g}_Z(\mathbf{B}) \equiv \text{vech}(\hat{\boldsymbol{\Omega}}) - \text{vech}(\mathbf{B}\mathbf{B}') = \mathbf{0},$$

where the operator $\text{vech}(\cdot)$ takes a symmetric $n \times n$ matrix and stacks the lower triangular half into a single vector of length $n\,(n+1)\,/2$. The VAR is therefore under-identified as there can be infinitely many solutions satisfying the covariance restrictions $\bar{g}_Z(\mathbf{B}) = \mathbf{0}$. Let these uncountably many solutions be collected into the set

$$\hat{\mathcal{B}} = \{\mathbf{B} = \hat{\mathbf{P}}\mathbf{Q} : \mathbf{Q} \in \mathbb{O}_n, diag\,(\mathbf{B}) \geq 0, \bar{g}_Z(\mathbf{B}) = \mathbf{0}\},$$

where $\mathbb{O}_n$ is the set of $n \times n$ orthonormal matrices. We shall refer to $\hat{\mathcal{B}}$ as the *unconstrained solution set* for short, with the understanding that it is not completely unconstrained given the imposition of the covariance restrictions. To simplify notation, the dependence of $\hat{\mathcal{B}}$ on $\mathbf{Q}$ and $\hat{\boldsymbol{\phi}}$ is suppressed.

To construct the unconstrained solution set $\hat{\mathcal{B}}$, we initialize $\mathbf{B}$ to be the unique lower-triangular Cholesky factor of $\hat{\boldsymbol{\Omega}}$ with non-negative diagonal elements, $\hat{\mathbf{P}}$, and then rotate it by

$K = 1.5$ million random orthogonal matrices $\mathbf{Q}$.[2] Each rotation begins by drawing an $n \times n$ matrix $\mathbf{M}$ of NID(0,1) random variables. Then $\mathbf{Q}$ is taken to be the orthonormal matrix in the $\mathbf{QR}$ decomposition of $\mathbf{M}$. Since $\mathbf{B} = \hat{\mathbf{P}}\mathbf{Q}$, the procedure imposes the covariance restrictions vech($\mathbf{\Omega}$) =vech($\mathbf{BB}'$) by construction. Let $\mathbf{e}_t(\mathbf{B}) = \mathbf{B}^{-1}\hat{\boldsymbol{\eta}}_t$ be the shocks implied by a $\mathbf{B} \in \hat{\mathcal{B}}$ for given $\hat{\boldsymbol{\eta}}_t$. The moments implied by the covariance structure alone give us 1.5 million values of $\mathbf{B}$, and thus 1.5 million unconstrained values of $\mathbf{e}_t(\mathbf{B})$ for $t = 1, ...T$.

Researchers have used various types of restrictions to dismiss solutions in $\hat{\mathcal{B}}$ leading to a smaller set that satisfies the additional identifying restrictions. A notable example is sign restrictions or more generally inequality restrictions of the form $\bar{g}_S(\mathbf{B}) \geq 0$. Existing theoretical and empirical work tends to place these constraints on the impulse response functions, i.e., in terms of $\mathbf{X}_t = \boldsymbol{\Psi}(L)\mathbf{B}\mathbf{e}_t$, the restrictions have focused on $\boldsymbol{\Psi}(L)\mathbf{B}$. Point identification requires restrictions beyond the ones implied by the covariance structure to reduce $\hat{\mathcal{B}}$ to a singleton.

2.1 Event Inequality Constraints

Event inequality constraints are unusual episodes of history in which a broad-based (historical and statistical) reading of the times would suggest a specific feature of the structural shocks. The idea is that a credible identification scheme should produce shocks that are not grossly at variance with our ex-post understanding of events, at least during periods of special interest. Event inequality constraints put bounds $\bar{\mathbf{k}} = \left(\bar{k}_1, ..., \bar{k}_E\right)$ on the sign and magnitude of $\mathbf{e}_t = \mathbf{B}^{-1}\boldsymbol{\eta}_t$ during selected episodes collected into a vector of E event dates $\bar{\boldsymbol{\tau}} = (\bar{\tau}_1, ..., \bar{\tau}_E)$. That is, each "event" is associated with a specific date or dates in the sample. These shocks are useful for identification because, from $\mathbf{e}_t = \mathbf{B}^{-1}\boldsymbol{\eta}_t = \mathbf{Q}'\mathbf{P}^{-1}\boldsymbol{\eta}_t$, we see that for $\tilde{\mathbf{Q}} \neq \mathbf{Q}$,

$$\tilde{\mathbf{e}}_t = \tilde{\mathbf{Q}}'\mathbf{P}^{-1}\boldsymbol{\eta}_t = \tilde{\mathbf{Q}}\mathbf{e}_t \neq \mathbf{e}_t$$

at any given t. This implies that constraints involving the shocks at specific time periods in the sample could be used to constrict the number of solutions in $\hat{\mathcal{B}}$. To illustrate the point, consider the $n = 2$ case:

$$\begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}.$$

so that solving for e_{1t} gives

$$e_{1t} = |\mathbf{B}|^{-1}(B_{22}\eta_{1t} - B_{12}\eta_{2t}),$$

where $|\mathbf{B}| = B_{11}B_{22} - B_{12}B_{21}$ is the determinant of $\mathbf{B}$. The values of $\eta_{1\tau}$ and $\eta_{2\tau}$ are given since we have data for event date τ in $\bar{\boldsymbol{\tau}}$. Hence, a restriction on the behavior of $e_{1\tau_1}$ at specific time $\bar{\tau}_1$ is a non-linear restriction on $\mathbf{B}$, or equivalently, on $\mathbf{Q}$. With non-Gaussian reduced-form errors, one can also see that the third and higher order moments of $\mathbf{e}_{1t}$ are not invariant to $\mathbf{B}$,

---

[2]The basic results discussed below are little affected by using larger numbers of draws, up to to 60 million.

hence $\mathbf{Q}$. This is in spite of the fact that the first and second moments of $\mathbf{e}_t$ are invariant to $\mathbf{Q}$. There is thus information in $\mathbf{e}_t$ that can be used to identify $\mathbf{B}$.

Several event inequality constraints may be represented as a system of inequality constraints on $\mathbf{B}$:

$$\bar{g}_E(\mathbf{e}_t(\mathbf{B}); \bar{\boldsymbol{\tau}}, \bar{\mathbf{k}}) \geq 0.$$

This is tantamount to creating dummy variables from the timing of specific events, and then putting restrictions on their correlation with the identified shocks. The motivation is that if a particular $\mathbf{Q}$ generates a shock series $\mathbf{e}_t$ that is difficult to defend in certain episodes, it can be removed from $\hat{\mathcal{B}}$. Such constraints could be imposed on extraordinary events such as the major recessions, wars, and natural disasters that have been well-documented. For example, if the first shock (say to monetary policy) is presumed to be strongly contractionary in $\bar{\boldsymbol{\tau}} = (1979{:}10, 1979{:}11, 1979{:}12)$, then one could formulate restrictions of the form

$$\bar{g}_E(\mathbf{B}; \bar{\boldsymbol{\tau}}, \bar{\mathbf{k}}) = \begin{pmatrix} -\sum_{t=1}^{T} 1_{t=1979{:}10} \cdot e_{1t} \\ -\sum_{t=1}^{T} 1_{t=1979{:}11} \cdot e_{1t} \\ -\sum_{t=1}^{T} 1_{t=1979{:}12} \cdot e_{1t} \end{pmatrix} - \begin{pmatrix} \bar{k}_1 \\ \bar{k}_1 \\ \bar{k}_1 \end{pmatrix} \geq 0$$

to dismiss solutions that imply highly expansionary monetary policy shocks in these episodes. The parameter $\bar{k}_1$ is a lower bound that reflects how contractionary these shocks are thought to be and represents a maintained assumption of the identification scheme, analogous to the zero restrictions of recursive identification schemes. The i-th row of $\bar{g}_{Ei}$ represents an inequality with $1_{t=\tau_i}$ as instrument. In essence, $\bar{g}_E(\mathbf{B}; \bar{\boldsymbol{\tau}}, \bar{\mathbf{k}})$ defines conditions based on the timing, sign, and magnitude of the events to help identification. Note that event inequality constraints put restrictions on the sign and the magnitude of $\mathbf{e}_t(\mathbf{B})$ rather than on the signs of impulse responses, as is common in some SVAR approaches.

Of course, if $\bar{k}_i$ is too big, or if the timing of the events in $\bar{\boldsymbol{\tau}}$ are inaccurate, the solutions will be meaningless even if they exist. On the other hand, if shocks are systematically found at particular episodes when no restrictions on the shocks are imposed, we can be more confident of their occurrence. Our approach is to construct the set of shocks in $\hat{\mathcal{B}}$ implied by the covariance structure alone, and then examine the properties of the shocks in the periods for which the $e_{it}(\mathbf{B})$ are large. This is exemplified in the applications below.

In concurrent work, Antolín-Díaz and Rubio-Ramírez (2018) also suggest using restrictions on the shocks during certain episodes of history to help identification in a Bayesian setting. They entertain event restrictions that play up the role of some shocks while simultaneously playing down the role of others (e.g., their "Type A and B" restrictions), similar in spirit to the traditional narrative-external instrument approach. Our event inequality constraints only require the weaker assumption that the events be driven at least in part by one or more of the shocks; they do not require the remaining shocks to play smaller roles. Furthermore, they do

not use external variables which play an important role in our approach, as we now discuss. A growing number of newer research studies use event restrictions of the general form suggested here.[3]

2.2 External Variable Inequality Constraints

When theory or economic reasoning imply that certain variables external to the VAR should be informative about the shocks of interest, such variables can also facilitate identification. Similar to the event inequality constraints, restrictions involving the correlation between $\mathbf{e}_t = \mathbf{B}^{-1}\boldsymbol{\eta}_t$ and external variables, can be used to constrict the number of solutions in $\hat{\mathcal{B}}$.

To be clear that external variables in this methodology are not necessarily valid instruments, we refer to the external variables presumed to have valuable information about the parameters of interest simply as $\mathbf{S}_t$. 'Valuable' is defined in terms of a lower bound on the unit-free correlation between $\mathbf{S}_t$ and the identified shocks, akin to the instrument relevance condition.

An example helps understand the motivation of these constraints. Consider a two variable model

$$\mathbf{A}(L)\begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}\begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}, \qquad \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix} \sim \mathcal{N}(0, \mathbf{I}_2),$$

where $\mathbf{I}_2$ is a 2×2 identity matrix. The covariance structure $\boldsymbol{\Omega} = \mathbf{B}\mathbf{B}'$ provides three unique pieces of information, so one more restriction would be needed for point identification. If an instrumental variable Z_t exists such that (i) $\mathbb{E}[Z_t e_{1t}] = 0$ (exogeneity) and (ii) $\mathbb{E}[Z_t e_{2t}] \neq 0$ (relevance), then a unique solution for $\mathbf{B}$ can be obtained.

Suppose that, instead of a valid *external instrument* Z_t, to use the terminology of Stock and Watson (2008), we have an external proxy variable S_t that is not assured to be exogenous and hence can be contemporaneously correlated with at least one structural shock. This suggests we could represent S_t by

$$S_t = \gamma e_{1t} + \Gamma e_{2t} + \sigma_S e_{St} \tag{1}$$

where e_{St} is an S-specific shock uncorrelated with e_{1t}, e_{2t} by assumption. We may want to discard solutions in $\hat{\mathcal{B}}$ for which the absolute correlation between S_t and e_{2t} is too small. The quantity

$$c(\mathbf{B}) = \frac{\Gamma}{\sqrt{\Gamma^2 + \gamma^2 + \sigma_S^2}},$$

measures the correlation between the component S_t and e_2. Requiring that $c(\mathbf{B}) > \bar{c}$ is the same as $\frac{\Gamma^2}{\Gamma^2 + \gamma^2 + \sigma_S^2} > \bar{c}^2$, which is a non-linear constraint on the parameters of the S_t equation, which are themselves functions of the shocks and data on S_t. That c is between zero and one facilitates the parameterization of $\bar{c}$.

---

[3]See for example, Cieslak and Pang (2019); Cieslak and Schrimpf (2019); Zeev (2018).

External variable restrictions can be collected into a system of inequality constraints on $\mathbf{B}$:

$$\bar{g}_C(\mathbf{e}_t(\mathbf{B}); \mathbf{S}, \bar{\mathbf{c}}) \geq 0,$$

where $\bar{\mathbf{c}}$ is a vector of lower bounds on the correlations between $e_t(\mathbf{B})$ and $\mathbf{S}$. In many applications $\bar{\mathbf{c}}$ can be close to but not exactly zero, thereby requiring that a shock merely be at least weakly correlated with $\mathbf{S}$. For example, if the first shock (say to monetary policy) is presumed to be correlated with an external variable S_t (say Fed fund futures), then one could formulate restrictions of the form

$$\bar{g}_C(\mathbf{e}_t(\mathbf{B}); \mathbf{S}, \bar{\mathbf{c}}) = corr\,(e_{1t}, S_t) - \bar{c} \geq 0.$$

Two points are worthy of emphasis. First, in conventional instrumental variable estimation, instrument exogeneity is a maintained assumption. By contrast, our approach makes no such exogeneity assumption. We only assert that the external variables be driven at least in part by one or more of the shocks, thereby allowing us to narrow the set of solutions but not achieve point identification. Of course, S_t itself is a valid exogenous instrument if $\gamma = 0$. But when validity of the exogeneity assumption is questionable, then S_t is at best *plausibly exogenous* in the terminology of Conley et al. (2012). These authors consider estimation of the equation for X_{1t} with endogenous regressor X_{2t} when instrument exogeneity is not known to hold exactly, but the parameter is point identified when the exogeneity assumption is valid. They put bounds on the effect due to the invalid instrument, or what we refer to as S_t, on X_{1t}. In contrast, we analyze X_{1t} and X_{2t} jointly, and we put more structure on the role of S_t so that a lower bound can be placed on its relevance for the shock or shocks of interest. Like Conley et al. (2012), such a bound will not, in general, be enough to achieve point identification. But it could dismiss solutions that do not achieve this bound, akin to dismissing weak instruments. Second, as in the external instrumental variable literature, a maintained assumption of the external variable inequality constraints is that the random processes behind the external variables are determined outside of the VAR system. External variables $\mathbf{S}_t$ are then well suited to help with identification whenever they are informative about the shocks of interest, while the research question is not concerned with their behavior *per se*. Otherwise they should be included in the SVAR system rather than used extraneously, which means they cannot be utilized to help with identification in the subsystem that excludes S_t.

2.3   Overview

In the applications below, the event and external variable inequality constraints are used individually or jointly, and possibly in conjunction with other types of restrictions, such as conventional sign restrictions on the IRFs. Suppose estimates of $\mathbf{B}$ are required to satisfy all such

restrictions for a given application. The *constrained solution set* is defined by

$$\bar{\mathcal{B}}(\mathbf{B}; \bar{\mathbf{k}}, \bar{\boldsymbol{\tau}}, \mathbf{S}) = \{\mathbf{B} = \hat{\mathbf{P}}\mathbf{Q} : \mathbf{Q} \in \mathbb{O}_n, \quad \text{diag}(\mathbf{B}) > 0;$$
$$\bar{g}_Z(\mathbf{B}) = 0, \bar{g}_S(\mathbf{B}) \geq 0, \bar{g}_E(\mathbf{e}_t(\mathbf{B}); \bar{\boldsymbol{\tau}}, \bar{\mathbf{k}}) \geq 0, \bar{g}_C(\mathbf{e}_t(\mathbf{B}); \mathbf{S}) \geq 0\}.$$

where $\bar{g}_Z(\mathbf{B}) = 0$ is the collection of covariance structure restrictions, $\bar{g}_S(\mathbf{B}) \geq 0$ is a set of sign and other restrictions on the IRFs, $\bar{g}_E(\mathbf{e}_t(\mathbf{B}); \bar{\boldsymbol{\tau}}, \bar{\mathbf{k}}) \geq 0$ is the set of event inequality constraints, and $\bar{g}_C(\mathbf{e}_t(\mathbf{B}); \mathbf{S}) \geq 0$ is the collection of external variable inequality constraints. To simplify notation, we simply write $\bar{\mathcal{B}}(\mathbf{B}; \bar{\mathbf{k}}, \bar{\boldsymbol{\tau}}, \bar{\lambda}, \mathbf{S})$ as $\bar{\mathcal{B}}$. A particular solution can be in both $\hat{\mathcal{B}}$ and $\bar{\mathcal{B}}$ only if all these restrictions are satisfied. Though $\bar{\mathcal{B}}$ is still a set, it should be smaller than $\hat{\mathcal{B}}$, which is based on the covariance restrictions alone. The additional identification restrictions that lead to $\bar{\mathcal{B}}$ explicitly recognize that not every solution in $\hat{\mathcal{B}}$ is equally credible.

Few methods are available to evaluate the sampling uncertainty of set identified SVARs from a frequentist perspective, and these tend to be specific to the imposition of particular identifying restrictions. Granziera, Moon and Schorfheide (2018) suggest a projections based method within a moment-inequality setup, but it is designed to study SVARs that only impose restrictions on one set of impulse response functions. Gafarov, Meier and Montiel-Olea (2015) suggest to collect parameters of the reduced-form model in a $1 - \alpha$ Wald ellipsoid but the approach is conservative. For the method to get an exact coverage of $1 - \alpha$, the radius of the Wald-ellipsoid needs to be carefully calibrated. As discussed in Kilian and Lutkepohl (2016), even with these adjustments, existing frequentist confidence sets for set-identified models still tend to be too wide to be informative. It is fair to say that there exists no generally agreed upon method for conducting inference in set-identified SVARs, let alone one proven to have correct frequentist coverage properties for shock-restricted SVARs of the type considered here.

We therefore use a bootstrap Monte Carlo procedure to assess the sampling error of our inequality restrictions when $\mathbf{S}_t$ are variables external to the SVAR. The Online Appendix describes a Monte Carlo simulation that bootstraps from the $\mathbf{e}_t(\mathbf{B})$ shocks for the $\mathbf{X}_t$ system to create confidence bands for impulse responses. These bootstrap standard errors are reported in several figures below.

We now turn to three SVAR studies to illustrate the potential usage of shock-based restrictions in particular applications.

3   Application 1: An Empirical Analysis of the Oil Market

In this section we consider an SVAR model of the oil market based on Kilian (2009) and Kilian and Murphy (2012) (KM hereafter). The objective of these studies is to determine the role of oil supply versus oil demand in driving volatility in oil price changes. These authors consider an SVAR with three variables: $\mathbf{X}_t = \begin{pmatrix} \Delta prod_t & rea_t & rpo_t \end{pmatrix}'$ where $\Delta prod_t$ is the percentage

change in global crude oil production, rea_t is the global demand of industrial commodities variable constructed in Kilian (2009), and rpo_t is the real oil price. Kilian (2009) refers to rea_t simply as a variable that measures "aggregate demand," for commodities in general, a concept to be distinguished from an oil market specific demand. The three structural shocks of interest to oil supply shock, aggregate demand shock, and to oil-specific demand shock. These are collected into the vector $\mathbf{e}_t$, which is related to the reduced-form errors η_t through

$$\begin{pmatrix} \eta_{\Delta prod,t} \\ \eta_{rea,t} \\ \eta_{rpo,t} \end{pmatrix} = \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{33} \\ B_{31} & B_{32} & B_{33} \end{pmatrix} \begin{pmatrix} e_{\Delta prod,t} \\ e_{rea,t} \\ e_{rpo,t} \end{pmatrix}. \tag{2}$$

3.1 Motivating Facts

Much has been written about the correlation between oil prices and geopolitical as well economic events. See, for example, Hamilton (2013) and Baumeister and Kilian (2016) for recent reviews. An extensive narrative history of big economic events in the oil market can also be found in Kilian and Murphy (2014) (KM2). However, their causal relations and more precisely the relative importance of the sources of fluctuations in the oil market is still a matter of debate.

Figure 1 plots the change in the real oil price over time in our sample, 1973:02 through 2004:09. The two largest spikes in the unconditional oil price change occur in (i) 1986:02, with a sharp downward spike following the collapse of OPEC, and (ii) 1990:08, with a sharp upward spike in the month of the Kuwait invasion by the U.S. If a reading of events in these months suggests that the large oil price changes are partly attributable to oil supply shocks, we would expect a spike of the appropriate sign in the structural $e_{\Delta prod,t}$ shock during these episodes. Therefore, we turn to the unconstrained set $\hat{\mathcal{B}}$ to see whether the data identify these as large oil supply shock events.

We first search over increasingly large numbers of rotations (up to 20 million) in $\hat{\mathcal{B}}$ over all time periods for the date with the smallest oil supply shock $e_{\Delta prod,t}$ (i.e., largest negative oil supply shock) and find that it is 1990:08, the month of the Kuwait invasion by the U.S. This also coincides with the date with the most minima in the oil supply shock $e_{\Delta prod,t}$ across all rotations. We further search over increasingly large numbers of rotations in $\hat{\mathcal{B}}$ over all time periods for the date with the biggest oil supply shock $e_{\Delta prod,t}$ (i.e., largest positive oil supply shock) and find that it is 1986:02 following the collapse of OPEC. This is also the date with the most maxima for $e_{\Delta prod,t}$ across all rotations. These findings are consistent with the hypothesis that the large spikes in the oil price during these months were at least partly attributable to oil supply disruptions. At the same time, KM2 have argued that the oil price increase in 1990:08 was caused by both a negative oil supply shock and a positive oil-specific demand shock. We therefore repeat the search over increasingly large numbers of rotations in $\hat{\mathcal{B}}$ over all time periods for the date with the largest oil specific demand shock $e_{rpo,t}$. Consistent with the arguments

10

Figure 1: Oil Price Change
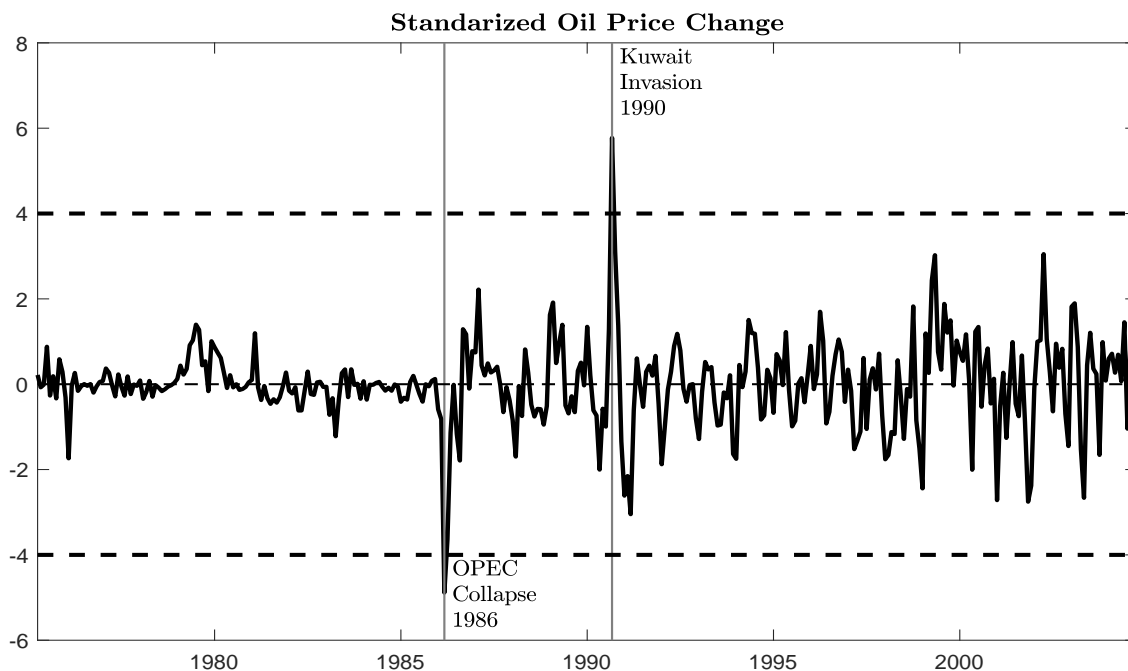
The figure plots the standardized change in the real oil price change. The sample spans the period 1975:01-2004:09.

in KM2, the date with the largest $e_{rpo,t}$ is again 1990:08, the month of the Kuwait invasion by the U.S. This is also the date in our sample with the most maxima in $e_{rpo,t}$ across all rotations. In addition, we find that the date with the smallest oil specific demand shock $e_{rpo,t}$ is 1986:02, following the collapse of OPEC, which is also the date with the most minima in $e_{rpo,t}$ across all rotations. Thus, the covariance structure of the data alone provides overwhelming evidence of both large oil supply and oil-specific demand shocks in the months 1990:08 and 1986:02, respectively. We note that the event dates identified using the above methodology remain the same regardless of the number of rotations, in a range from 1.5 million to 20 million. We therefore use the smaller number for the estimation discussed below.

As observed by KM2, the collapse of OPEC in early 1986 was preceded by an announcement in late 1985 by Saudi Arabia that it would no longer attempt to prop up the price of oil by reducing its oil production. According to the historical evidence presented in KM2, the actual supply disruption that created a major positive shock to the flow supply of oil around the OPEC collapse took place over several months between 1985:12 and roughly 1986:06, and was accompanied in the short run by lower storage demand as oil price expectations fell. The historical account given in Hamilton (2013) generally agrees and argues that the Saudi's "ramped up" production around this time, leading to a positive "oil supply shock for producers." This suggests that the OPEC collapse may be better characterized as a sequence of positive oil

supply shocks between 1985:12 and 1986:06, rather than a single shock in one month.

Kilian (2008) provides an "exogenous oil supply shock" external variable series that measures shortfalls in OPEC oil production associated with wars and civil disruptions. This indicator is used as an external *instrument* for point identifying oil price shocks in Montiel-Olea, Stock and Watson (2015) (MSW). Production shortfalls would be a valid instrumental variable if it were uncorrelated with the two demand shocks $e_{rea,t}$ and $e_{rpo,t}$. The assumption was used in MSW to point identify oil supply shocks $e_{\Delta prod,t}$ in the above SVAR. But the variable could be an imperfect indicator of actual production shortfalls and measurement error is possible. Furthermore, some have argued that oil production shortfalls are at least partly caused by political events such as wars or embargoes that could have direct implications for oil demand (e.g., Hamilton (2013)). Yet, even if the shortfall series is not truly exogenous with respect to the two demand shocks, it may still be relevant for oil supply shocks. Our approach is to use the information in the shortfall variable by exploiting its correlation with the oil supply shock but without insisting on zero correlations with the two oil demand shocks.

3.2 Shock-Based Constraints

With this background in hand, we now consider several different combinations of restrictions to aid with identification. These are summarized in Table 1 below. The row labeled $\bar{g}_Z = 0$ denotes a weak set of constraints based on the covariance restrictions alone. Kilian (2009) combines the covariance restrictions with traditional sign restrictions on the impact impulse response functions. This combination of restrictions is referred to as the K09 restrictions in Table 1. KM combine the K09 restrictions with restrictions on certain combinations of parameters in the **B** matrix that place upper bounds on the ratios $\frac{B_{13}}{B_{33}}$ and $\frac{B_{12}}{B_{32}}$.[4] Results from these combined restrictions are used to form a basis of comparison with our distinct shock-based restrictions and are labeled KKM in Table 1.

Next we consider adding both event inequality constraints $\bar{g}_E \geq 0$ and external variable inequality constraints $\bar{g}_C \geq 0$ to help with identification of oil supply versus aggregate and oil demand shocks. These are combined with the covariance and K09 sign restrictions and labeled SEE in Table 1. Event inequality constraints are summarized by $\bar{g}_{E_1}$ and $\bar{g}_{E_2}$ in row SEE of the table. Constraint $\bar{g}_{E1}$ requires both a large negative oil supply shock and a large positive oil-specific demand shock in the month of the Kuwait invasion. Specifically, constraint $\bar{g}_{E_1}$ requires that the $e_{\Delta prod,\bar{\tau}_1}$ found in period $\bar{\tau}_1$ of August 1990 be small and less than $\bar{k}_1$ standard deviations below the mean and that the $e_{rpo,\bar{\tau}_1}$ be large and exceed $\bar{k}_2$ standard deviations above

---

[4]KM interpret the restrictions on the ratios $\frac{B_{13}}{B_{33}}$ and $\frac{B_{12}}{B_{32}}$ as restrictions on *elasticities* of oil supply to demand shocks. There is, however, disagreement about whether these particular restrictions isolate the relevant elasticities (see the debate between Baumeister and Hamilton (2019) and Kilian (2019)). We take no stand on the interpretation of these restrictions and instead merely implement the same restrictions used in KM solely to form a basis of comparison with the shock-based restrictions of this paper.

Table 1: Identification Restrictions, Oil Application

Model	Restrictions	Notes
$\bar{g}_Z = 0$	$\text{vec}(\hat{\Omega}) - \text{vech}(BB') = 0$	covariance restrictions
K09	$\bar{g}_Z = 0$ $B_{11} < 0, B_{12} > 0, B_{13} > 0$ $B_{21} < 0, B_{22} > 0, B_{23} < 0$ $B_{31} > 0, B_{32} > 0, B_{33} > 0$	covariance restrictions sign restrictions (Kilian (2009))
KKM	$\bar{g}_Z = 0$, K09, $\frac{B_{13}}{B_{33}} < 0.258$, $\frac{B_{12}}{B_{32}} < 0.258$	covariance, sign, and KM restrictions
SEE	$\bar{g}_Z = 0$ K09 $\bar{g}_{E_1}$: $\bar{k}_1 - e_{\Delta prod, \bar{\tau}_1} \geq 0, e_{rpo, \bar{\tau}_1} - \bar{k}_2 \geq 0$ $\bar{g}_{E_2}$: $\sum_{\tau \in \bar{\tau}_2} e_{\Delta prod, \tau} \geq 0, -\sum_{\tau \in \bar{\tau}_2} e_{rpo, \tau} \geq 0$ $\bar{g}_C$: $0 \geq \text{corr}(OS_t, e_{\Delta prod, t})$	covariance restrictions sign restrictions $\bar{\tau}_1 = 1990{:}08$ (Kuwait invasion) $\bar{k}_1 = -2.4$ std. $= q_{50}(e_{\Delta prod, \bar{\tau}_1})$ $\bar{k}_2 = 2.9$ std. $= q_{50}(e_{rpo.\bar{\tau}_1})$ $\bar{\tau}_2 = [1985{:}12, 1986{:}06]$ (OPEC collapse) $S_t = OS_t$: oil production shortfall series
SEE(a)		$\bar{k}_1 = -4.5$ std. $= q_{25}(e_{\Delta prod, \bar{\tau}_1})$ $\bar{k}_2 = 2.9$ std. $= q_{50}(e_{rpo, \bar{\tau}_1})$
SEE(b)		$\bar{k}_1 = -0.3$ std. $= q_{75}(e_{\Delta prod, \bar{\tau}_1})$ $\bar{k}_2 = 2.9$ std. $= q_{50}(e_{rpo, \bar{\tau}_1})$
SEE(c)		$\bar{k}_1 = -2.4$ std. $= q_{50}(e_{\Delta prod, \bar{\tau}_1})$ $\bar{k}_2 = 4.5$ std. $= q_{75}(e_{rpo, \bar{\tau}_1})$
SEE(d)		$\bar{k}_1 = -2.4$ std. $= q_{50}(e_{\Delta prod, \bar{\tau}_1})$ $\bar{k}_2 = 0.7$ std. $= q_{25}(e_{rpo, \bar{\tau}_1})$
MSW	$\bar{g}_Z = 0$ $\mathbb{E}[Z_t e_{\Delta prod, t}] \neq 0$ $\mathbb{E}[Z_t e_{rea, t}] = \mathbb{E}[Z_t e_{rpo, t}] = 0$	covariance restrictions $Z_t = OS_t$ used as external instrument

Notes: The table summarizes restrictions in different constrained solution sets. $q_\alpha(x)$ refers to the α percentile value of x.

the mean. Constraint $\bar{g}_{E_2}$ requires both positive oil supply and negative oil-specific demand shocks during the OPEC collapse. Specifically, the constraint requires the cumulation of $e_{\Delta prod, t}$ in $\bar{\tau}_2 = [1985{:}12, 1986{:}06]$ to be non-negative, which is to say that their sum must be above average, and the cumulation of $e_{rpo, t}$ in $\bar{\tau}_2 = [1985{:}12, 1986{:}06]$ to be non-positive, which is to say that their sum may not be above average. The restrictions labeled SEE and SEE(a)-SEE(d) set different values for the threshold parameters $\bar{\mathbf{k}} = (\bar{k}_1, \bar{k}_2)'$ and will be discussed below.

Note that there is nothing in the event inequality constraints that explicitly precludes the presence of the global demand shock $e_{rea, t}$ from playing an important role in these episodes. Nor is there anything that restricts the relative importance of the three shocks in any particular episode. The event inequality constraints merely require that oil supply and oil-specific shocks played some role in the price spikes of these episodes, where the magnitude and sign of that

role are determined by the parameters $\bar{k}_1$ and $\bar{k}_2$.

The external variable inequality constraints are summarized as $\bar{g}_C \geq 0$ in row SEE of Table 1. We use the "oil production shortfall" series constructed by Kilian (2008) and denoted OS_t as our S_t.[5] The constraint requires that any oil supply shocks $e_{\Delta prod,t}(\mathbf{B})$ formed from $\mathbf{B} \in \hat{\mathcal{B}}$ be negatively correlated with OS_t, implying that production shortfalls resulting from wars and embargoes coincide with decreases in $e_{\Delta prod,t}$, or negative oil supply shocks.

Finally, the restrictions labeled MSW describe the constraints used in MSW to point identify oil supply shocks $e_{\Delta prod,t}$ in the above SVAR. These restrictions assume OS_t is a valid external instrument and we therefore label it Z_t in that case. MSW are concerned that OS_t could be a weak instrument. By contrast, the SEE constraints do not treat OS_t as a valid instrument, but nonetheless assume that it is relevant for supply shocks.

3.3   Results for the Oil SVAR

We use the same data used in Kilian (2009) and the largest common sample period across Kilian (2008) and Kilian (2009) (1973:02-2004:09) for the analysis. Following KM, we use $p = 24$ lags in the VAR to capture the long swings in the oil market. After losing observations to lags and differencing, the estimation sample is 1975:02 to 2004:09. We study the dynamic causal effects and propagating mechanisms of the shocks under different constraints and parameterizations using impulse response functions. All figures show responses to one standard deviation shocks in the direction that raise the price of oil.

Figure 2 shows the IRFs of the real price of oil to an oil supply shock, an aggregate commodity demand shock, and an oil market specific demand shock, under several different combinations of identifying restrictions. The IRFs using the KM restrictions are shown in black dotted lines of the left panel of Figure 2. As emphasized by KM and KM2, the bounds of these sets are wide and display a substantial range of responses to all three shocks. Among 1.5 million rotations, $4,878$ solutions satisfy both the covariance and the sign restrictions. For comparison, the top two panels plot the IRF obtained by the point-identified model of MSW, which uses OS_t as an external instrument to recover $e_{\Delta prod,t}$. This estimate implies that oil supply shocks have tiny if not zero effects on oil prices. Indeed they are smaller for several months after the shock than that implied by even the lower bound of the solution set obtained under K09 restriction.

Next we consider the IRFs using the KKM restrictions These are shown in grey shaded areas on the left panel of Figure 2. Only 34 solutions satisfy this combination of constraints, implying that the KM constraints alone severely shrink the number of admissible solutions. The dynamic responses to movements in $e_{\Delta prod,t}$ under the KKM restrictions are numerically small (top panel), indicating that the solutions generating the large responses of oil price to a supply

---

[5]The data for OS_t are based on the replication files of Montiel-Olea et al. (2015) and span the period 1973:02 to 2004:09.

Figure 2: Impulse Response of Real Oil Prices

Sign and KM Restrictions

Sign + Event + External Variable Inequality Constraints
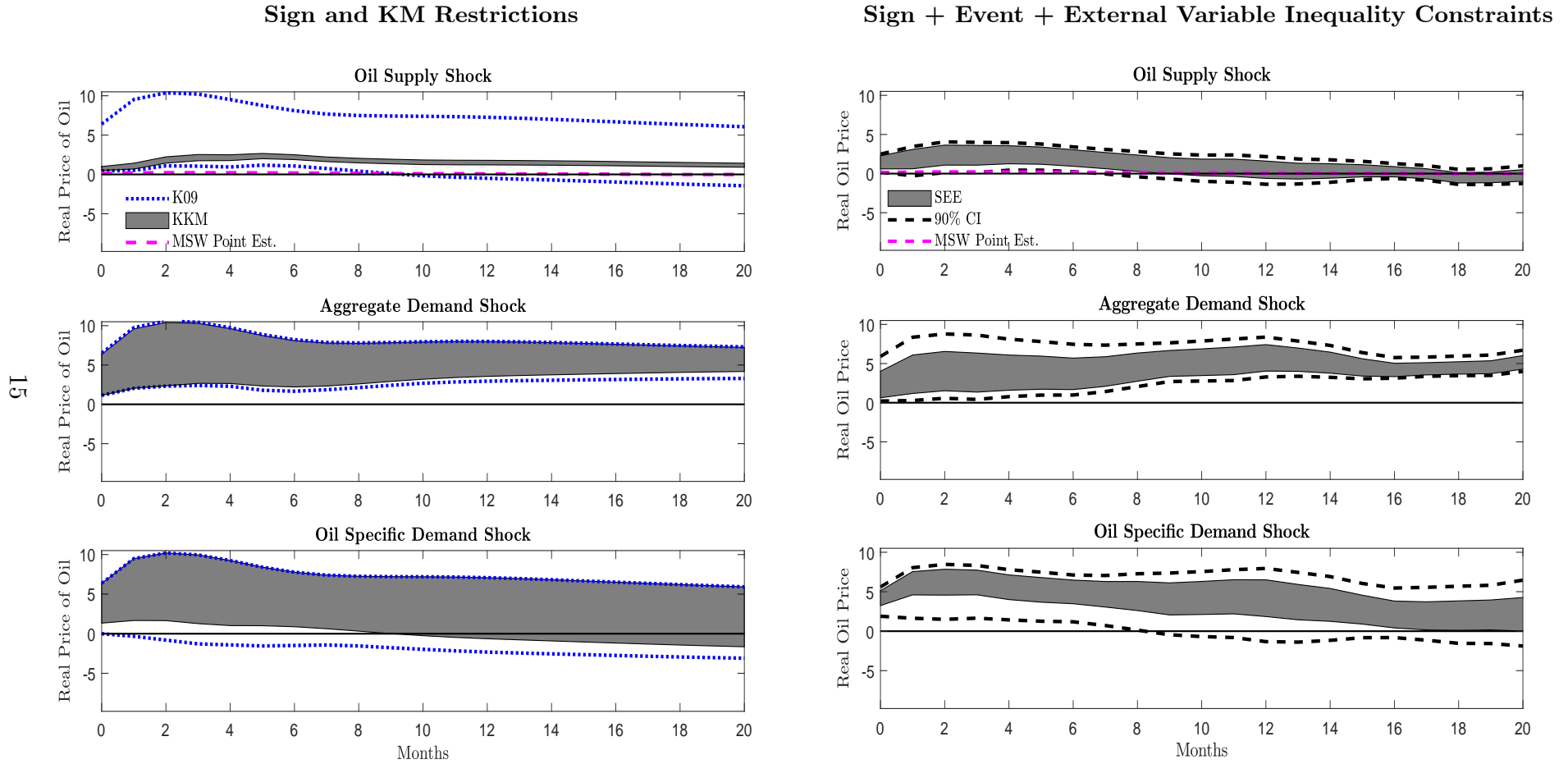
The figure reports solution sets of impulse response to positive, one standard deviation shocks for system $\mathbf{X}_t = (\Delta prod_t, rea_t, rpo_t)'$ under KM, KKM, MSW and SEE restrictions listed in Table 1. The sample spans the period 1973:02 to 2004:09.

shock have been eliminated by the KM restrictions on $\frac{B_{13}}{B_{33}}$ and $\frac{B_{12}}{B_{32}}$. Instead, the aggregate demand shock $e_{rea,t}$ has large effects on the price of oil, both in the short- and long-run. Some KKM solutions also imply that the oil-specific demand shock $e_{rpo,t}$ has quantitatively large effects on oil price changes, but the upper and lower bound of the responses to this shock are much wider and thus the findings in this regard are less conclusive.

We now consider the SEE restrictions which replace the KM restrictions with, sign, event, and external variable inequality constraints. To implement these restrictions, we need to set parameters of the "big shock" event inequality constraints $\overline{\mathbf{k}} = \left(\overline{k}_1, \overline{k}_2\right)'$. To do so, we return to the unconstrained set $\hat{\mathcal{B}}$ to examine the observed magnitudes of the shocks in these episodes. We first examine a very large number of rotations (20 million) to locate the largest absolute values of $e_{\Delta prod,\overline{\tau}_1}$ and $e_{rpo,\overline{\tau}_1}$ in $\overline{\tau}_1 = 1990{:}08$ and find that the largest negative value for $e_{\Delta prod,\overline{\tau}_1}$ is -5.98 standard deviations, while the largest positive value for $e_{rpo,\overline{\tau}_1}$ is 5.98 standard deviations. The threshold parameters of the big shock constraints set lower bounds for the absolute magnitude of the shocks in these episodes, thus we choose them so that the relevant constraint is considerably less restrictive than what would be implied by these largest absolute values. To start, we set $\overline{\mathbf{k}} = \left(\overline{k}_1, \overline{k}_2\right)' = (-2.4, 2.9)$. These numbers can be more readily interpreted by noting that they would roughly correspond to the 50th percentile values of $e_{\Delta prod,\overline{\tau}_1}$ and $e_{rpo,\overline{\tau}_1}$ in $\hat{\mathcal{B}}$ in $\overline{\tau}_1 = 1990{:}08$. The constraints therefore require a "big" negative oil supply shock to be *at most* 2.4 standard deviations below the mean (in the bottom 50% of all $e_{\Delta prod,\overline{\tau}_1}$ in $\hat{\mathcal{B}}$ in 1990:08), and they require a "big" positive oil-specific demand shock to be *at least* 2.9 standard deviations above the mean (in the top 50% of all $e_{rpo,\overline{\tau}_1}$). The values for $\overline{\mathbf{k}}$ are reported in Table 1 along with the corresponding percentile values. We denote the α percentile value of x as $q_\alpha(x)$.

We also check that the absolute magnitude of the shocks on these event dates remains for all practical purposes the same regardless of the number of rotations, in a range from 1.5 million to 20 million. For example, the smallest $e_{rpo,\overline{\tau}_1}$ in $\hat{\mathcal{B}}$ is -5.97 when examining 1.5 million draws, and is -5.98 for 20 million. We therefore use the smaller number of draws for the various estimations discussed below.

The IRFs associated with the SEE restrictions under the above parameterization of $\overline{\mathbf{k}}$ are shown as grey shaded areas in the right panel of Figure 2, along with 90 percent confidence bands computed using a boot-strap sampling approach. Comparing results across the two panels, we find that the set of responses to the two demand shocks are tighter than those under the KKM restrictions, and preserve solutions that produce larger oil price responses to demand shocks of either type than to supply shocks. These results are similar qualitatively to those reported in KM. One difference between the sets produced by the SEE constraints versus the KKM restrictions is that the range of responses to the oil supply shock under the SEE restrictions includes values that are somewhat larger than under the KKM restrictions, even though the effects of oil supply shocks are still smaller than those of the demand shocks. The

16

MSW point estimates give responses to oil supply shocks that fall outside the bounds of both the KKM and SEE solution sets.

All identification approaches require assumptions, and it is important to check the sensitivity of the results to model assumptions. For the present application, the parameters $\overline{\mathbf{k}} = (\bar{k}_1, \bar{k}_2)'$ stipulate lower bounds in absolute terms for the oil supply or oil-specific demand shocks, respectively, during the Kuwait invasion. To assess the sensitivity of the results the this parameterization, Figure 3 shows the IRFs under four different values for these parameters. Case SEE(a) sets $\bar{k}_1$ to -4.5 standard deviations, which is the $q_{25}(e_1)$ percentile value of $e_{\Delta prod,t}$ in $\hat{\mathcal{B}}$ at $\bar{\tau}_1$ while keeping $\bar{k}_2$ at $q_{50}(e_{rpo,t})$. Case SEE(b) sets $\bar{k}_1$ to -0.3 standard deviations, which is the $q_{75}(e_1)$ percentile value of $e_{\Delta prod,t}$ in $\hat{\mathcal{B}}$ at $\bar{\tau}_1$, also keeping $\bar{k}_2$ while keeping $\bar{k}_2$ at $q_{50}(e_{rpo,t})$. The first parameterization tightens while the second loosens the first event inequality constraint relative to the SEE case. Case SEE(c) keeps $\bar{k}_1$ at $q_{50}(e_{\Delta prod,t})$ but sets $\bar{k}_2$ to 4.5 standard deviations, which is the 75th percentile value of $e_{rpo,t}$ in $\hat{\mathcal{B}}$ at $\bar{\tau}_2$, while Case SEE(d) keeps $\bar{k}_1$ at $q_{50}(e_{\Delta prod,t})$ and sets $\bar{k}_2$ to 0.7 standard deviations, which corresponds to the 25th percentile value of $e_{rpo,t}$ in $\hat{\mathcal{B}}$ at $\bar{\tau}_2$. Case SEE(c) loosens while SEE(d) tightens the second event inequality constraint relative to the SEE case. The left panel of Figure 3 shows the IRFs for SEE(a) and SEE(b) while the right panel shows the IRFs for SEE(c) and SEE(d). Its clear from the figure that the qualitative results are not sensitive to the parameterization of the the the parameterization of the event inequality constraint. Under all four parameterizations, positive shocks to both types of demand lead to a sharp increases in the price of oil that persists for many months, while the effects of oil supply shocks are more muted. No matter which parameterization, the responses to the aggregate demand shock is bounded well away from zero as the horizon increases.

3.4 Properties of the Shocks

Although a stated objective of any SVAR analysis is to identify the structural shocks, the properties of the shocks are rarely scrutinized. By contrast, in the methodology here the shocks are of explicit interest, so we examine their properties. One property of interest concerns the normality of the shocks. The $e_{\Delta prod,t}$ identified by imposing the SEE restrictions exhibit strong non-Gaussian features. Averaged across solutions, the coefficient of skewness and kurtosis are -0.6102 and 5.4865, respectively. But the $e_{\Delta prod,t}$ series implied by the KKM restrictions exhibit even greater departures from normality, with an average skewness of -1.4603 and a kurtosis of 11.0722. The stronger departures from Gaussianity arises because the KKM constraints accept solutions that imply larger $e_{\Delta prod,t}$ shocks in some time periods.

To have a clearer picture of the properties of the shocks identified by the SEE versus KKM restrictions, Figure 4 plots the timing of "large shocks" in both cases, where for the SEE restriction case we set $\overline{\mathbf{k}} = (\bar{k}_1, \bar{k}_2)'$ equal to the 50th percentile values in $\hat{\mathcal{B}}$ of the relevant

## Figure 3: IRFs under Different Parameterizations of Big Shocks: Oil Application

Weaker/tighter constraints on $e_{\Delta prod}$ in 1990:08

Weaker/tighter constraints on e_{rpo} in 1990:08
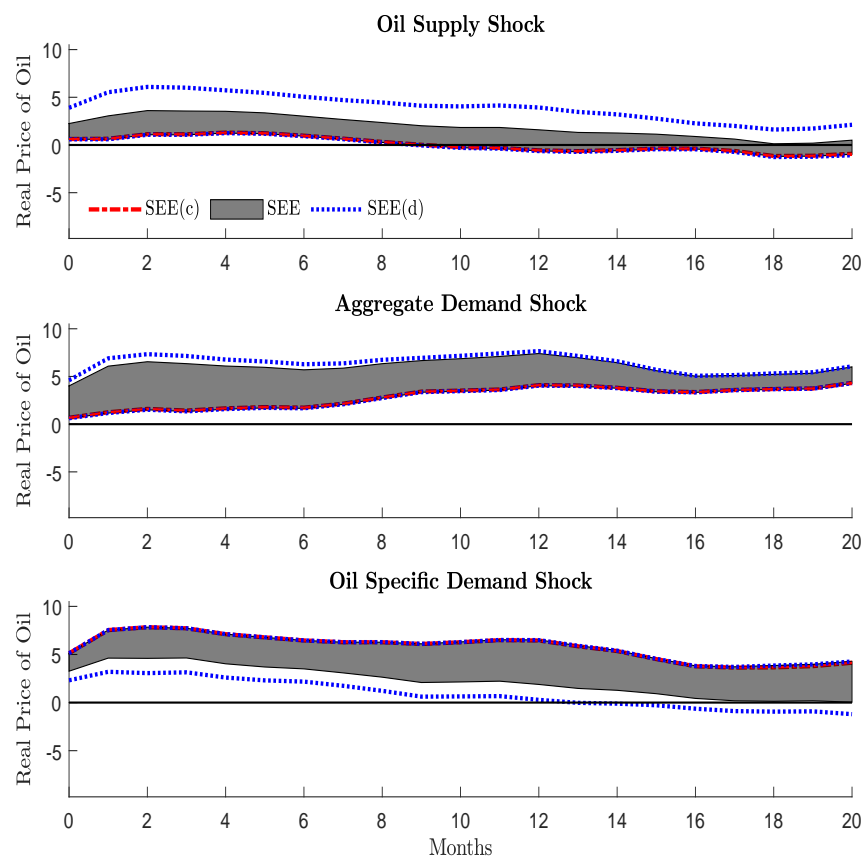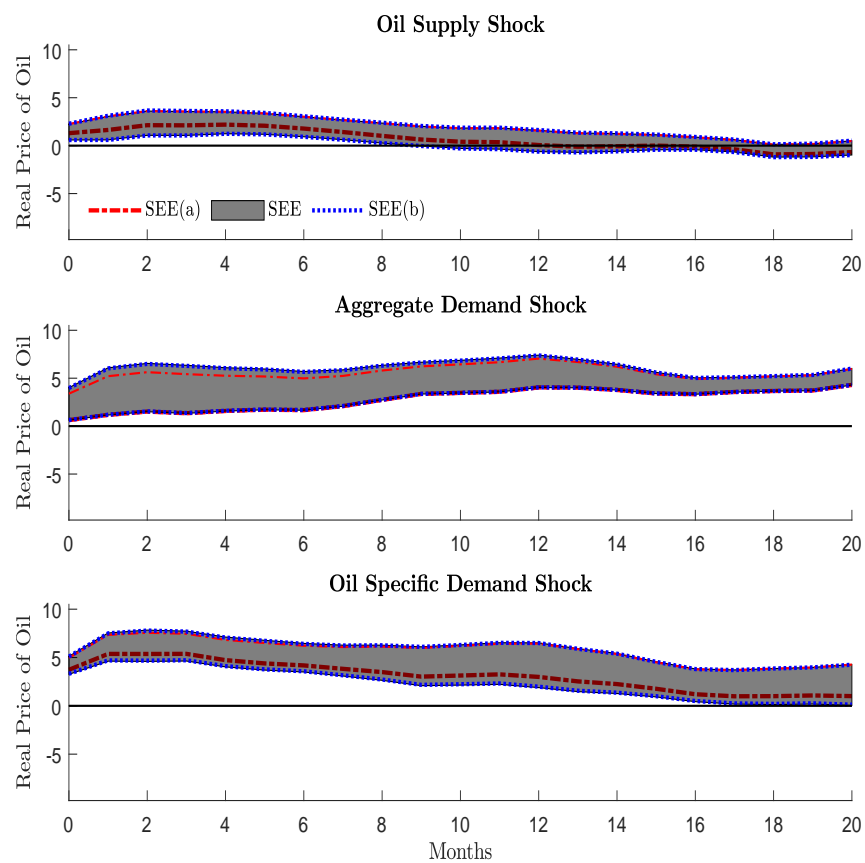
The figure reports solution set of impulse response to positive, one standard deviation shocks for system $\mathbf{X}_t = (\Delta prod_t, rea_t, rpo_t)'$ under SEE, SEE(a) to SEE (d) restrictions listed in Table 1. SEE (a) and SEE (c) correspond to tighter event inequality constraints whereas SEE(b) and SEE(d) corresponds to weaker event inequality constraints. The sample spans the period 1973:02 to 2004:09.

Figure 4: Large Shocks in Oil Application
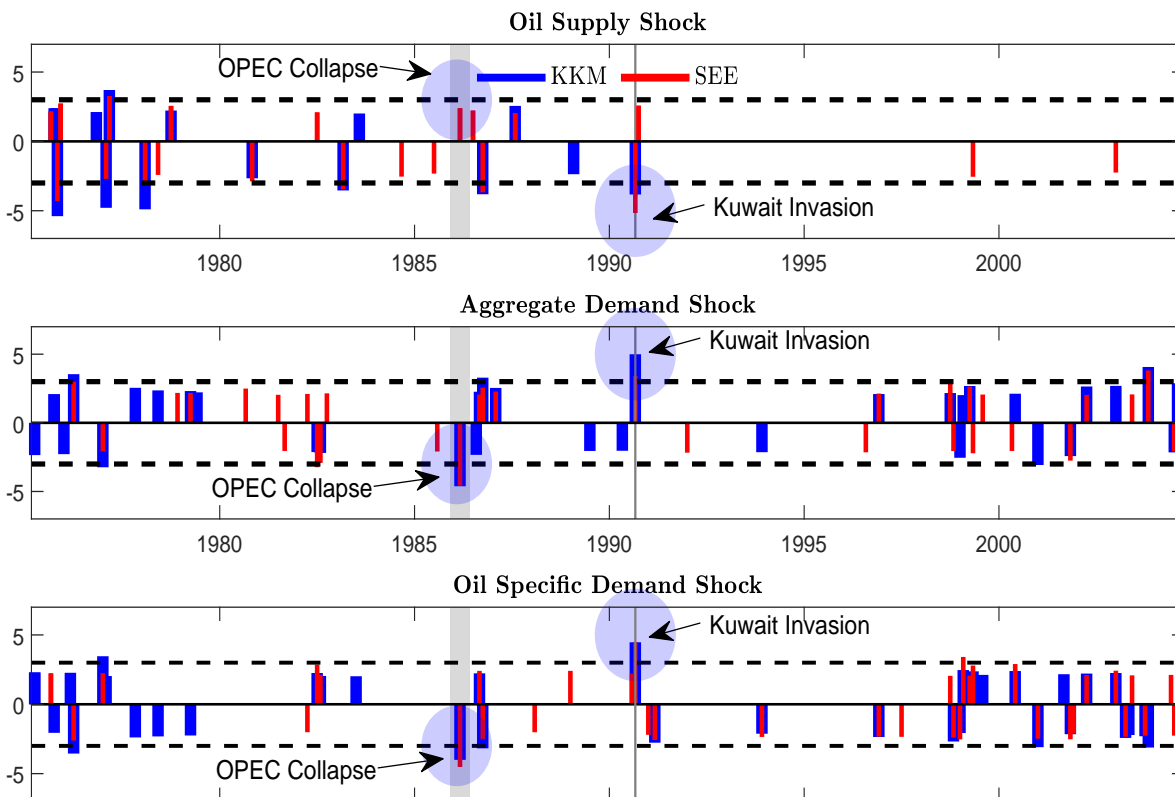
The figure shows all shocks in the solution set that are at least 2 standard deviations above/below the unconditional mean from the solution set for system $\mathbf{X}_t = (\Delta prod_t, rea_t, rpo_t)'$ under KKM and SEE restrictions listed in Table 1. The thin vertical line shows the date 1990:08 and the shaded vertical bar shows the range of dates associated with the OPEC collapse, 1985:12-1986:06. The horizontal line corresponds to 3 standard deviations above/below the unconditional mean of each series. The sample spans the period 1973:02 to 2004:09.

shocks in 1990:08. For the purposes of the figure, large shocks are defined to be those in excess of two standard deviations above (or below) the mean. In view of the non-normality of the shocks, the figure also plots horizontal lines corresponding to three standard deviation of the unit shocks. The figure reports the standard deviation of all such large shocks in the identified sets $\bar{\mathcal{B}}$ based on the SEE restrictions and compares them to those under the KKM restrictions.[6] By design, the $e_{\Delta prod,t}$ shocks generated by the SEE restrictions, displayed in red, should be less than -2.45 standard deviations in 1990:08, corresponding to the median value in $\hat{\mathcal{B}}$ in 1990:08. In fact, on that date, 49 solutions have a supply shock less than -5 standard deviations. Moreover, this is the only date in the solution sets that exhibit values for $e_{\Delta prod,t}$ that are smaller than negative five standard deviations. By way of comparison, the KKM restrictions also identify large negative supply shocks in the month of the Kuwait invasion.

---

[6]If multiple solutions have shocks on a given date that are larger in absolute value than two standard

On the other hand, the spotlighted area in Figure 4 around the OPEC collapse shows that the KKM restrictions produce no big positive oil supply shocks at any time during this episode. This implication of the KKM restrictions appears empirically implausible on the basis of a broadly-shared ex post understanding of the OPEC collapse. Indeed, both KM2 and Hamilton (2013) agree that Saudi Arabia created a major positive "shock" to the flow supply of oil between the end of 1985 and the middle of 1986, which contributed to a large drop in its price. These findings underscore how identification schemes derived exclusively from attention to parameter restrictions may miss valuable clues from the data that can help evaluate the validity of the identifying restrictions.

Figure 4 shows that the SEE constraints by no means preclude big $e_{rea,t}$ shocks from occurring in 1990:08 and $\overline{\tau}_2 \in [1985:12, 1986:06]$, even though behavior of $e_{rea,t}$ was left unconstrained throughout the entire sample. Moreover, the SEE constraints generate several large $e_{rea,t}$ and $e_{rpo,t}$ of both signs outside of our events, and large $e_{\Delta prod,t}$ outside of constrained event dates are numerous under both the KKM and SEE identification schemes. Regardless of which identification scheme is used, large negative supply shocks coexist with large positive demand shocks in 1990:08.

We close this section by investigating possible reasons why the MSW point estimate diverges so much from the set-identified results under the SEE restrictions. MSW use the oil shortfall series OS_t as an external instrument for identifying $e_{\Delta prod,t}$, which explicitly imposes the exogeneity assumption that OS_t be uncorrelated with both $e_{rpo,t}$ and $e_{rea,t}$. The solution sets given by the SEE restrictions impose no such assumption but are free to recover it if they are consistent with those restrictions. Figure 5 presents a histogram showing the distribution of estimated correlations between the OS_t and $e_{rea,t}$ (top panel), and between OS_t and $e_{rpo,t}$ (bottom panel) across all solutions in the constrained solution set $\bar{\mathcal{B}}$ under the SEE parameter values for $\overline{\mathbf{k}} = (\overline{k}_1, \overline{k}_2)'$. While the 90 percent confidence bands for the correlations between OS_t and $e_{rea,t}$ include zero, there is no solution that exhibits a zero correlation between OS_t and $e_{rpo,t}$. Given that exogeneity of OS_t is inconsistent with the SEE restrictions, the MSW point estimate can only lie outside the bounds of the solution set, thereby explaining the divergent results.

4   Application 2: Monetary Policy and Financial Markets

In this section we consider an SVAR application based on the baseline specification in Gertler and Karadi (2015) (GK hereafter), which studies role of monetary policy shocks on the aggregate economy. The GK system is comprised of the following variables: $\mathbf{X}_t = (i_t, cpi_t, ip_t, ebp_t)'$,

---

deviations, they are represented in the by bars that lie exactly on top of one another. Thus, the length of the bar indicates the size of the *largest* shock for the month, among all solutions.

Figure 5: Distribution of Correlations with Oil Demand Shocks
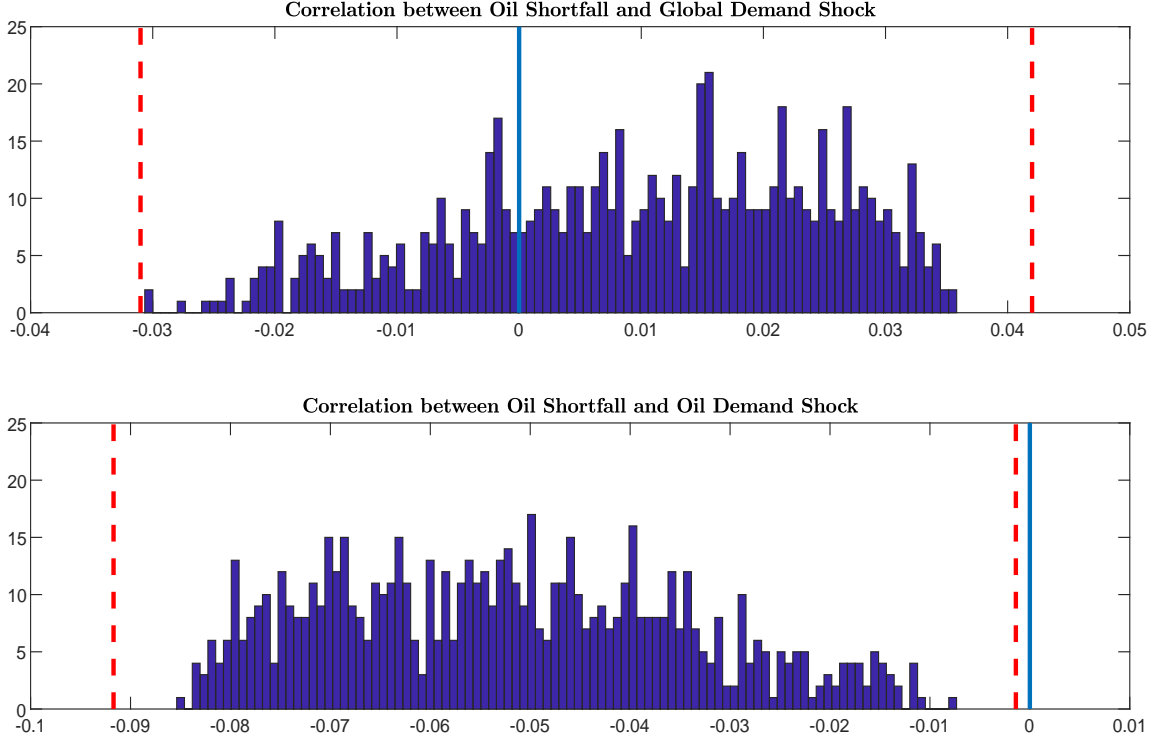
The figure displays histograms for all values of correlations between oil shortfall and global demand shock (top panel) and correlations between oil shortfall and oil specific demand shock (bottom panel) in an solution set for system $\mathbf{X}_t = (\Delta prod_t, rea_t, rpo_t)'$ under SEE restriction listed in Table 1. The sample spans the period 1973:02 to 2004:09.

where i_t is a "policy indicator," measured here as the one-year Treasury bill (t-bill) rate, cpi_t is the log of the Consumer Price Index, ip_t is the log of industrial production, and ebp_t is the excess bond premium (EBP) of Gilchrist and Zakrajšek (2012), a measure of credit spreads.

Let the four structural shocks of the SVAR be collected into the vector

$$\mathbf{e}_t = (e_{i,t}, e_{cpi,t}, e_{ip,t}, e_{ebp,t})'.$$

The first shock, $e_{i,t}$, is the monetary policy shock of interest. This shock is the component of the VAR forecast error for i_t that is uncorrelated with the other structural shocks in the SVAR. We refer to $e_{cpi,t}$, $e_{ip,t}$, and $e_{ebp,t}$ as "non-policy" shocks.

The objective in GK is to identify $e_{i,t}$ and trace out its dynamic effects on financial market variables such as ebp_t. To do so, they construct the difference in the price of the 3-month fed funds futures contract between 20 minutes after and 10 minutes before a Federal Open Market Committee (FOMC) announcement. Since these are surprise movements in the fed funds futures price in tight windows of monetary policy announcements, they are arguably attributable to monetary policy. The VAR requires monthly data, so GK turn the futures surprises on FOMC

days into monthly average surprises by allocating them between consecutive calendar months based on when they happened within the calendar month.[7] This monthly variable (denoted FF4 in GK) is the external instrument and will be denoted Z_t below. GK then point identify the policy shock $e_{i,t}$ by maintaining the following assumptions:

$$\mathbb{E}\left[Z_t e_{i,t}\right] = \phi \neq 0 \ \ \text{(inst. relevance)} \tag{3}$$

$$\mathbb{E}[Z_t e_{cpi,t}] = \mathbb{E}[Z_t e_{ip,t}] = \mathbb{E}[Z_t e_{ebp,t}] = 0 \ \text{(inst. exogeneity)} \tag{4}$$

Restriction (3) requires Z_t to have a non-zero correlation with the policy shock it identifies, implying that the instrument must be relevant. Restriction (4) requires Z_t to be uncorrelated with all the other shocks in the system, implying that the instrument must be exogenous.

At this point it is useful to note that the term exogenous has different meanings in different contexts. It is reasonable to assume, as GK do, that innovations in the fed funds futures prices on FOMC days capture "exogenous" movements in Federal Reserve policy in the sense that they are not influenced by macroeconomic and financial conditions around tight windows of FOMC announcements. The restrictions in (4), however, require a type of exogeneity with a different connotation. Specifically, they require that the innovations in fed funds futures prices around FOMC announcements in a month have no contemporaneous influence on the forecast errors of the non-policy variables, except insofar as they affect $e_{i,t}$.

While the assumptions inherent in (4) are a reasonable starting place, relaxing them is of interest for at least two reasons. First, a number of authors have argued that financial variables other than short-term Treasury rates serve as additional policy indicators that capture distinct channels of monetary policy transmission (e.g., Gagnon, Raskin, Remache and Sack (2011), Krishnamurthy and Vissing-Jorgensen (2011), Krishnamurthy and Vissing-Jorgensen (2013), Swanson (2017)). If true, the exogeneity assumption (4) for variables such as ebp_t may be overly strong, since credit spreads are likely to reflect disparate policy transmission channels not captured by short-term interest rates. Second, and relatedly, the time-aggregation required to turn futures market surprises on FOMC days into monthly average surprises makes the exogeneity assumptions less tenable. For example, an FOMC announcement that occurs in the middle of the month must be assumed to have no relation with the non-policy shocks $e_{cpi,t}, e_{ip,t}, e_{ebp,t}$ at any time within that month, including the two weeks following the announcement. If monetary policy has effects that operate through channels orthogonal to short-term interest rates, e.g., if they affect other assets which in turn quickly affect investment plans and/or credit spreads, the exogeneity restriction (4) would be invalid.

According to either reason, (4) could be relaxed. The first reason suggests that variables such as ebp_t should be included in a vector of policy indicators along with i_t. In this case,

---

[7]Specifically, if they happened at the beginning of the month, GK allocate most of the surprises to the month (and the rest to the next month). If they happened at the end of the month, GK allocate them mostly to the next month. We refer the reader to GK for further details.

there would no longer be a single policy shock $e_{i,t}$ in the system but instead a vector of policy shocks $(e_{i,t}, e_{ebp,t})'$, while the scalar parameter ϕ would become a two dimensional vector $\boldsymbol{\phi}$. The second reason suggests that any shock $e_{cpi,t}, e_{ip,t}, e_{ebp,t}$ could possibly be correlated with Z_t contemporaneously. In either case, the external instrumental variable approach to point estimation can no longer be implemented because, with only a single instrument Z_t, the system is now under-identified. While this renders the external instrument approach inoperative, it is not a problem for the shock-restricted approach, which obviates the need for the exogeneity restrictions. The challenge with the shock-restricted approach is to find other credible identifying restrictions capable of substantively winnowing the number of solutions in $\hat{\mathcal{B}}$. In the next section we illustrate how different types of shock-based restrictions can be used to generate solution sets that still give a fairly clear picture of the dynamic causal effects of $e_{i,t}$ shocks in the GK system.

4.1   Motivating Facts

Since we are interested in understanding monetary policy shocks that affect short-term interest rates, it is useful to isolate historical episodes characterized by quantitatively large monetary policy effects on these rates. To do so, we turn to a well-known historical record of such shocks that is based on the Greenbook residual series presented in Romer and Romer (2004). The Greenbook residuals show changes in the "intended" federal funds rate not taken in response to Federal Reserve Greenbook forecasts about inflation or real growth, formed by taking the residuals from a multivariate regression of the intended funds rate on the Greenbook forecasts. A negative (positive) residual indicates a negative (positive) monetary policy surprise. To locate quantitatively large policy surprises, we isolate dates for which these Greenbook residuals were either greater than the 95th percentile value of their sample observations, or less than the 5th percentile value. We refer to the dates that satisfy these criteria as the *GB95 dates* (big tightenings) and the *GB05 dates* (big easings), respectively.

Next we investigate whether these dates correspond to "big shock" events using the covariance structure restrictions alone. We again search over increasingly large numbers of rotations (up to 20 million) in $\hat{\mathcal{B}}$ over the GB05 dates for the date with the smallest $e_{i,t}$ (i.e., the biggest surprise easing) and find that it is 1981:10. This also coincides with the date in our sample with the most minima of $e_{i,t}$ across all rotations. The date with the second smallest $e_{i,t}$ (i.e., the second biggest surprise easing) also falls on the date 1981:10. We therefore search for the next smallest $e_{i,t}$ that occurs on a *unique* date among the GB05 dates and find that it is 2001:11. This is also the date in our sample with the second most minima of $e_{i,t}$ across all rotations among the GB05 dates. (In the full sample this is the date with the third most minima.) Likewise, we search over increasingly large numbers of rotations (up to 20 million) in $\hat{\mathcal{B}}$ in the GB95

period for the date with the largest $e_{i,t}$ (i.e., the biggest surprise tightening) and find that it is 1981:05. This also coincides with the date in our sample that has the most maxima of $e_{i,t}$ across all rotations among the GB95 dates. (In the full sample this is the date with the third most maxima.) The date with the second largest $e_{i,t}$ (i.e., the second biggest surprise tightening) is also 1981:05. The next largest $e_{i,t}$ on a *unique* date in the GB95 dates is 1987:05. This is also the date in our sample with the second most maxima of $e_{i,t}$ across all rotations among the GB95 dates. As for the previous application, we check that the event dates identified using the above methodology remain the same regardless of the number of rotations, in a range from 1.5 million to 20 million. We therefore use the smaller number for the estimation discussed below.

Figure 6 plots the time series of i_t (top panel), the Greenbook residuals (middle panel), and all $e_{i,t}$ shocks in $\hat{\mathcal{B}}$ that are at least two standard deviations above or below the mean. The horizontal lines in the middle panel indicate the 95th and 5th percentile values of the Greenbook residuals. The horizontal lines in the bottom panel indicate three standard deviations above or below the mean for $e_{i,t}$. What we see is that the date 1981:10 has both a very negative Greenbook residual and one of the most negative observations on $e_{i,t}$ in the solution set over the entire sample (close to -5 standard deviations). While some solutions in $\hat{\mathcal{B}}$ in the post 2005 period have $e_{i,t}$ smaller than -5, relatively few solutions have this property relative to the dates 1981:10 and 2001:11.[8] Likewise, 1981:05 is both one of the largest tightening episodes according to the Greenbook residuals, and a very large tightening episode according to the unconstrained set of $e_{i,t}$ in that month.

Other major economic events in our sample, even if they are not big monetary policy events, may be valid candidates for restricting the behavior of the shocks. In October of 2008, following the collapse of Lehman brothers in late September of 2008, the Dow Jones Industrial average began a pronounced decline and subsequently fell more than 50% over a period of 17 months. The collapse in the market over this period has been associated with a broad-based financial crisis that is often cited as a "trigger" of the Great Recession. We argue that these dates are likely to be associated with a sharp increase in credit spreads, thereby justifying restrictions on the behavior of $e_{ebt,t}$ during these months of the financial crisis.

4.2 Shock-Based Constraints

Motivated by the historical facts just discussed, we now consider two types of shock-based restrictions summarized in the Table below. For ease of reference, the first row re-stipulates the restrictions used in GK and labels the "GK restrictions."

We use the fed funds futures surprise FF4 series as S_t for this application. The constraints

---

[8]If multiple solutions have $e_{i,t}$ on a given date that are larger in absolute value than two standard deviations, they are represented in the bottom panel by bars that lie exactly on top of one another. Thus, the length of the bar indicates the size of the largest shock for the month, among all solutions.

Figure 6: Monetary Variables
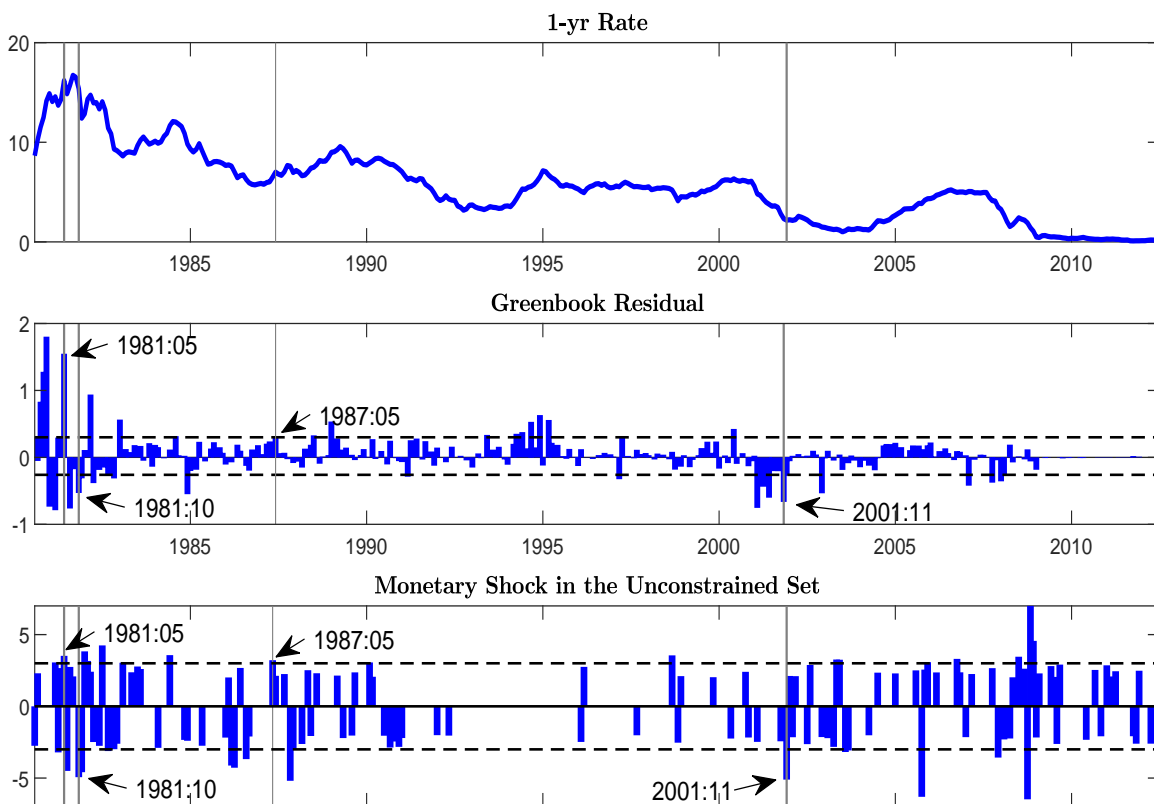
The figure plots the time series of one-year tbill rate $i_t^{(12)}$, Romer and Romer (2004) Greenbook residuals, and all shocks that are at least 2 standard deviations above/below the unconditional mean from the unconstrained set for system $\mathbf{X}_t = \left(i_t^{(12)}, cpi_t, ip_t, ebp_t \right)'$. The horizontal line in the middle panel corresponds to the 5th and 95h percentile of Greenbook residuals. The horizontal line in the bottom panel corresponds to 3 standard deviations above/below the unconditional mean. The thin vertical line corresponds to the dates chosen for the event inequality constraints. The data span the period 1979:07 to 2012:06.

labeled S1 employ two external variable inequality constraints $\bar{g}_{C_1} \geq 0$ and $\bar{g}_{C_2} \geq 0$ that require $S_t = \text{FF4}_t$ to be positively correlated with both $e_{i,t}$ and $e_{ebp,t}$. Though we do not treat FF4_t as a valid instrument, it must be at least weakly correlated with both shocks to address the concern that monetary policy might operate through multiple financial indicators and channels. The constraints under S2 use all the constraints in S1 and add two more external variable constraints, labeled $\bar{g}_{C3}$ and $\bar{g}_{C4}$. One would expect FF4_t to be approximately uncorrelated contemporaneously with the shocks of "slower moving" macro variables, even if such an assumption is unrealistic for financial market variables such as the EBP. Thus constraints $\bar{g}_{C3}$ and $\bar{g}_{C4}$ require that the contemporaneous relation between $e_{cpi,t}$ and S_t, and between $e_{ip,t}$ and S_t to be very small, though not numerically zero, implying that FF4_t is assumed to be approximately exogenous with respect to these variables. The results we present below are robust with values of ϵ as large as about 0.008. We present results for much smaller values.

Table 2: Identification Restrictions, Monetary Application

Model	Restrictions	Notes
$\bar{g}_Z = 0$	$\text{vec}(\hat{\Omega}) - \text{vech}(BB') = 0$	covariance restrictions
GK	$\bar{g}_Z = 0$ $\mathbb{E}[Z_t e_{i,t}] \neq 0$ $\mathbb{E}[Z_t e_{cpi,t}] = \mathbb{E}[Z_t e_{ip,t}] = \mathbb{E}[Z_t e_{ebp,t}] = 0$	covariance restrictions $Z_t = \text{FF4}_t$
S1	$\bar{g}_Z = 0$ $\bar{g}_{C1}$: $\text{corr}(S_t, e_{i,t}) \geq 0$ $\bar{g}_{C2}$: $\text{corr}(S_t, e_{ebp,t}) \geq 0$	covariance restrictions $S_t = \text{FF4}_t$
S2	S1 $\bar{g}_{C3}$: $|\text{corr}(S_t, e_{cpi,t})| \leq \epsilon$ $\bar{g}_{C4}$: $|\text{corr}(S_t, e_{ip,t})| \leq \epsilon$	$\epsilon = 10^{-4}$
SEE	S1 $\bar{g}_{E1} : (e_{ebp,\bar{\tau}_1} \geq \bar{k}_1) \vee (e_{ebp,\bar{\tau}_2} \geq \bar{k}_2)$ $\bar{g}_{E2} : e_{i,\bar{\tau}_3} \geq \bar{k}_3$ and $e_{i,\bar{\tau}_4} \geq \bar{k}_4$ $\bar{g}_{E3} : e_{i,\bar{\tau}_5} \leq \bar{k}_5$ and $e_{i,\bar{\tau}_6} \leq \bar{k}_6$	$\bar{\tau}_1$=2008:09, $\bar{\tau}_2$= 2008:10 (Lehman Collapse) $\bar{k}_1 = 3.4$ std $= q_{75}(e_{ebp,\bar{\tau}_1})$ $\bar{k}_2 = 5.4$ std $= q_{75}(e_{edp,\bar{\tau}_2})$ $\bar{\tau}_3$=1981:05, $\bar{\tau}_4$=1987:05 (Monetary Tightenings) $\bar{k}_3 = 2.6$ std. $= q_{75}(e_{i,\bar{\tau}_3})$ $\bar{k}_4 = 2.3$ std. $= q_{75}(e_{i,\bar{\tau}_4})$ $\bar{\tau}_5$=1981:10, $\bar{\tau}_6$=2001:11 (Monetary Easings) $\bar{k}_5 = -3.5$ std. $= q_{25}(e_{i,\bar{\tau}_5})$ $\bar{k}_6 = -2.0$ std. $= q_{25}(e_{i,\bar{\tau}_6})$
SEE(a)		$\bar{k}_3 = 1.7$ std. $= q_{50}(e_{i,\bar{\tau}_3})$ $\bar{k}_4 = 0.7$ std. $= q_{50}(e_{i,\bar{\tau}_4})$ $\bar{k}_5 = -3.5$ std. $= q_{25}(e_{i,\bar{\tau}_5})$ $\bar{k}_6 = -2.0$ std. $= q_{25}(e_{i,\bar{\tau}_6})$
SEE(b)		$\bar{k}_3 = 2.6$ std. $= q_{75}(e_{i,\bar{\tau}_3})$ $\bar{k}_4 = 2.3$ std. $= q_{75}(e_{i,\bar{\tau}_4})$ $\bar{k}_5 = -2.5$ std. $= q_{50}(e_{i,\bar{\tau}_5})$ $\bar{k}_6 = -0.4$ std. $= q_{50}(e_{i,\bar{\tau}_6})$

Notes: The table summarizes restrictions in different constrained solution sets. $q_\alpha(x)$ refers to the α percentile value of x.

The models denoted SEE, SEE(a), and SEE(b) add event inequality constraints, but do away with constraints $\bar{g}_{C3}$ and $\bar{g}_{C4}$ that restrict the correlations between FF4_t and $e_{cpi,t}$ and $e_{ip,t}$ to both be close to zero. SEE uses the constraints in S1 and adds $\bar{g}_{E_1}$, $\bar{g}_{E_2}$, and $\bar{g}_{E_3}$. Event inequality constraint $\bar{g}_{E_1} \geq 0$ requires that the EBP credit spread shock be large and exceed $\bar{k}_1$ and $\bar{k}_2$ standard deviations above the mean in either $\bar{\tau}_1 = 2008:09$ or $\bar{\tau}_2 = 2008:10$ (or both), for these months associated with the Lehman collapse and its immediate aftermath. Event inequality constraint $g_{E_2} \geq 0$ requires that the monetary policy shock exceed $\bar{k}_3$ and $\bar{k}_4$ standard deviations above the mean during months $\bar{\tau}_3 = 1981:05$ and $\bar{\tau}_4 = 1987:05$, respectively, when the Greenbook residuals were unusually large and positive. Event inequality constraint $g_{E_3} \geq 0$

requires that the monetary policy shocks be smaller than $\bar{k}_5$ and $\bar{k}_6$ standard deviations below the mean during months $\bar{\tau}_5 = 1981{:}10$ and $\bar{\tau}_6 = 2001{:}11$, respectively, when the Greenbook residuals were unusually small. The alternatives listed in SEE(a) and SEE(b) use different parameterizations of the parameters $\bar{k}_3$ - $\bar{k}_6$ that control how large a "big" monetary policy shock must be and will be discussed below.

4.3 Results for the Monetary Policy SVAR

We follow GK and set the VAR lag to be 12 months. Where available, we use the same monthly data used by GK and otherwise stick to their sample dates, which spans the period 1979:07 to 2012:06.[9] Although the VAR variables are available over this sample, the external variable $\mathrm{FF4}_t$ is available only from 1991:01 onward. We therefore follow GK and use the full sample 1979:07 to 2012:06 to estimate the reduced-form VAR, but impose $\bar{g}_{C1}$–$\bar{g}_{C4}$ only for the post 1991 subsample. We focus on the dynamic response of the EBP to a monetary policy shock, the main object of interest in GK. All figures show IRFs to a one standard deviation change in $e_{i,t}$ in the direction that raises i_t. For reference, the figures also display the GK point estimated IRFs.

To assess how the different constraints affect the identified impulse response functions, we begin with the impulse responses under restrictions SC1 which allow both $e_{ebp,t}$ and $e_{i,t}$ to be correlated with S_t, but do away with all of the exogeneity restrictions. The IRFs, plotted in the upper left panel of Figure 7, evidently are wholly uninformative and include both positive and negative responses of credit spreads to an impulse in $e_{i,t}$.

We now employ additional constraints to shrink the set. The first case is SC2, which adds $\bar{g}_{C3}$ and $\bar{g}_{C4}$ to SC1 to allow $e_{cpi,t}$ and $e_{ip,t}$ but not $e_{ebp,t}$ to have small correlation with S_t. The IRFs using the SC2 restrictions are shown as gray shaded areas in the upper left panel of Figure 7. The responses of this set are now highly informative as all responses in the set show that a positive impulse to $e_{i,t}$ drives up the EBP sharply. These results are similar to those produced by the GK point estimate, but obtained under weaker restrictions that permit $e_{ebp,t}$ to be contemporaneously correlated with $S_t = \mathrm{FF4}_t$.

From the last result, it is evident that constraints $\bar{g}_{C3}$ and $\bar{g}_{C4}$ have substantial identifying power. But they require that $S_t = \mathrm{FF4}_t$ to have close to zero contemporaneous relation with the structural shocks to inflation and production. We now eliminate these restrictions and instead replace them with the event inequality constraints $\bar{g}_{E1}, \bar{g}_{E2}$ and $\bar{g}_{E3}$. To implement these SEE constraints, we need to set the parameters of our big shock events $\bar{\mathbf{k}} = \left(\bar{k}_1, ..., \bar{k}_6\right)'$. We examine a large number of rotations (up 20 million) to locate the largest values of $e_{ebp,t}$ in absolute terms

---

[9]We replicated the original Romer and Romer (2004) Greenbook residuals over their sample 1979:07 to 1996:12 and then extended the residuals to the end of the GK sample, 2012:06. A more detailed description of Romer and Romer (2004) methodology is provided in the Online Appendix.

Figure 7: Impulse Response to Monetary Shock
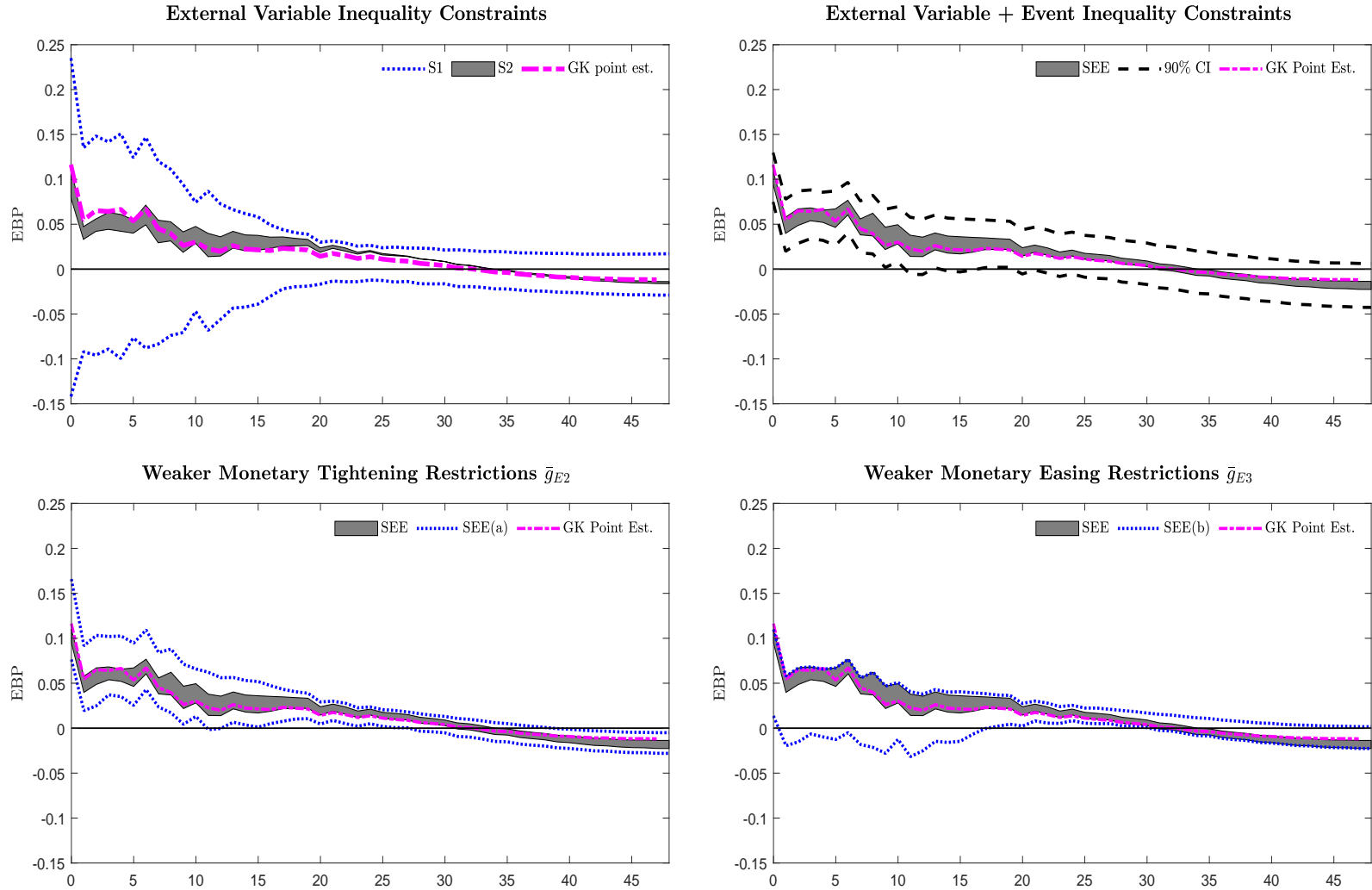
The figure reports solution set of impulse response to positive, one standard deviation shocks for system $\mathbf{X}_t = (i_t, cpi_t, ip_t, ebp_t)'$ under restrictions S1, S2, GK, SEE, SEE(a) and SEE(b) listed in Table 2. The sample spans the period 1979:07 to 2012:06.

in $t \in \{\bar{\tau}_1, \bar{\tau}_2\}$, and find that the largest value for $e_{ebp,\bar{\tau}_1}$ equals 6.48 standard deviations above the mean, while the largest value for $e_{ebp,\bar{\tau}_2}$ equals 7.33 standard deviations above the mean. Similarly, we examine a large number of rotations (up 20 million) to locate the largest values of $e_{i,t}$ in absolute terms in $t \in \{\bar{\tau}_3, \bar{\tau}_4, \bar{\tau}_5, \bar{\tau}_6\}$, and find that the largest value for e_{i,τ_3} is 3.52 standard deviations above the mean, the largest value for $e_{i,\bar{\tau}_4}$ is 3.13 standard deviations above the mean, the largest negative value for $e_{i,\bar{\tau}_5}$ is -4.94 standard deviations below the mean, and the largest negative value for $e_{i,\bar{\tau}_6}$ is -5.08 standard deviations below the mean. Parameters $\bar{\mathbf{k}}$ of the big shock event inequality constraints stipulate lower bounds in absolute terms for the shocks in these episodes, and thus should be considerably less restrictive than these largest absolute values across all rotations. For the EBP event restriction, $\bar{g}_{E1}$, we start by setting $(\bar{k}_1, \bar{k}_2)' = (3.5, 5.4)'$, which may be more readily interpreted by noting that these values would correspond to the the 75th percentile values of $e_{ebp,t}$ in $\hat{\mathcal{B}}$ for $t \in \{\bar{\tau}_1, \bar{\tau}_2\}$. Recall that event constraint $\bar{g}_{E1}$ requires that the EBP credit spread shock be large and exceed $\bar{k}_1$ and $\bar{k}_2$ standard deviations above the mean in either $\bar{\tau}_1 = 2008{:}09$ or $\bar{\tau}_2 = 2008{:}10$ (or both). For the monetary tightening event constraints, $\bar{g}_{E2}$, we start by setting $(\bar{k}_3, \bar{k}_4)' = (2.6, 2.3)'$, which may be more readily interpreted by noting that these values correspond to the the 75th percentile values of policy shock $e_{i,t}$ in $\hat{\mathcal{B}}$ in $t \in \{\bar{\tau}_3, \bar{\tau}_4\}$. Recall that the event constraint $\bar{g}_{E2}$ requires that the policy shock be large and exceed $\bar{k}_3$ and $\bar{k}_4$ standard deviations above the mean in $\bar{\tau}_3 = 1981{:}05$ and $\bar{\tau}_4 = 1987{:}05$. Thus this parameterization requires that a "big" interest rate shock must be in the top 75% of all shocks in $\hat{\mathcal{B}}$ during events when there were big positive Greenbook residuals. The event constraint $\bar{g}_{E3}$ hat the policy shock be sufficiently negative and less than or equal to the threshold parameters $(\bar{k}_5, \bar{k}_6)'$. We start by setting $(\bar{k}_5, \bar{k}_6)' = (-3.5, -2.0)'$, which may be more readily interpreted by noting that these values would correspond to 25th percentile values of $e_{i,t}$ in $\hat{\mathcal{B}}$ in $\bar{\tau}_5 = 1981{:}10$ and $\bar{\tau}_6 = 2001{:}11$ when there were big negative Greenbook residuals. As for the previous application, we check that the absolute magnitude of the shocks on these event dates remains for all practical purposes the same regardless of the number of rotations in a range from 1.5 million to 20 million. We therefore use the smaller number of draws for the various estimations discussed below.

The upper right panel of Figure 7 shows the IRFs for dynamic responses under the SEE constraints. Only 234 solutions satisfy these constraints but they are highly informative and show that a positive impulse in $e_{i,t}$ sharply increases credit spreads, as measured by the EBP, which remain elevated for many months. These results are qualitatively similar to those produced by the GK point estimate, lending support to the GK identifying restrictions.

We now consider the sensitivity of the results to the parameters $\bar{k}$. Case SEE(a) sets $(\bar{k}_3, \bar{k}_4) = (1.7, 0.7)'$, so that the big shocks can be smaller than the ones in SEE. The constraints are therefore less restrictive and threshold values correspond to the median rather than 75th percentile values of $e_{i,t}$ at $\bar{\tau}_3, \bar{\tau}_4$ in $\hat{\mathcal{B}}$. Case SEE(b) similarly sets $(\bar{k}_5, \bar{k}_6)' = (-2.5, -0.4)'$, so

that the big negative shocks at these dates may be less negative than the ones under the SEE restrictions. The constraints are therefore less restrictive and threshold values correspond to the median rather than 75th percentile values of $e_{i,t}$ at $\bar{\tau}_5, \bar{\tau}_6$ in in $\hat{\mathcal{B}}$. The bottom left panel of Figure 7 shows the IRFs under SEE(a) while the bottom right panel of Figure 7 shows the IRFs under SEE(b). In both cases, $\bar{k}_1$ and $\bar{k}_2$ are held fixed at the 75th percentile values as in SEE.

While sets of IRFs are inevitably wider when the constraints are weakened, the left panel shows that all solutions under weaker constraints for the tightening events still indicate that the EBP rises sharply in response to a monetary tightening. By contrast, the bottom right panel shows that when we weaken the event inequality constraints pertaining to monetary policy easings, the causal effect of a monetary policy tightening on credit spreads becomes inconclusive due to the wide range of solutions retained. Specifically, the responses of the EBP in the bottom right panel contain both positive and negative values for many periods after the shock. This finding suggests that monetary policy easings are more important for identification of monetary policy shocks than are monetary policy tightenings.

Is this asymmetry specific to the shock-restricted approach? To address this question we return to the GK external instrumental variable point-identification scheme, but instead of using $Z_t = \text{FF4}_t$ as an instrument, as in GK, we consider two alternative instruments. In the first case we set $Z_t = \text{FF4}_t \cdot I\left(\Delta\text{FF4}_t{<}0\right)$, where $I\left(\Delta\text{FF4}_t{<}0\right)$ is an indicator variable that equals one if the surprise movement in the futures rate on FOMC days is negative (i.e., a surprise easing), and zero otherwise. In the second case we set $Z_t = \text{FF4}_t \cdot I\left(\Delta\text{FF4}_t{>}0\right)$, where $I\left(\Delta\text{FF4}_t{>}0\right)$ is an indicator variable that equals one if the surprise movements in the futures rate on FOMC days is positive (i.e., a surprise tightening), and zero otherwise. Figure 8 shows the IRFs obtained under these identification schemes, along with the GK point estimate that uses $Z_t = \text{FF4}_t$. It is immediately evident that the IRFs obtained using $Z_t = \text{FF4}_t \cdot I\left(\Delta\text{FF4}_t{<}0\right)$ are almost identical to those produced with $Z_t = \text{FF4}_t$. By contrast, the results using $Z_t = \text{FF4}_t \cdot I\left(\Delta\text{FF4}_t{>}0\right)$ differ dramatically and imply that a positive monetary policy shock has a negligible impact on the EBP at all horizons and moreover has the "wrong" (negative) sign in the short-run. Thus much like the shock-restricted sets of IRFs in the bottom panels of Figure 7, the identifying power of the restrictions in this point-identified case appears largely attributable to monetary *easings*, with monetary tightenings playing little role.[10]

---

[10]Results of a first-stage F test to assess instrument strength for i_t show that $Z_t = \text{FF4}_t$ delivers an F statistic that passes the threshold value of 10 dictated by Stock, Wright and Yogo (2002) for a "non-weak" instrument. But this result too is entirely attributable to the monetary easings. The variable $Z_t = \text{FF4}_t \cdot I\left(\Delta\text{FF4}_t{<}0\right)$ delivers an F statistic equal to 19.78, while $Z_t = \text{FF4}_t \cdot I\left(\Delta\text{FF4}_t{>}0\right)$ generates an F statistic equal to 2.

Figure 8: Identification using Tightenings v.s. Easings: Alternative IV
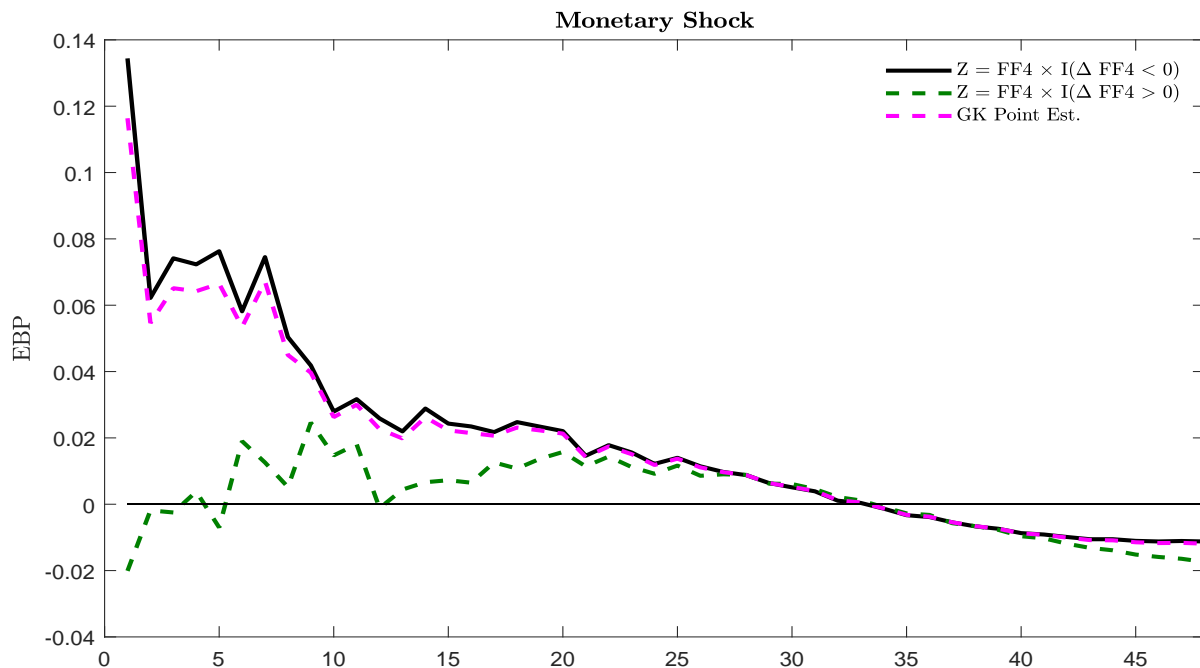
This figure reports the impulse response function for the GK point-identification approach with different instruments. The pink dotted line corresponds to the GK point estimate. The black (green) line corresponds to results using FF4 as an instrument over the subsample when FF4 decreases (increases). The sample spans the period 1979:07 to 2012:06.

5 Application 3: Housing Wealth and Consumption

As a final application we consider a VAR inspired by the work of Mian and Sufi (2014) to estimate the macroeconomic effects of housing wealth shocks on aggregate consumption. Mian and Sufi considered the role of house price gains in U.S. spending from 2002-2006, a time of rapid home price appreciation. Their data consisted of a panel of zip-code level observations, an ideal dataset for estimating how the consumption effects of housing wealth might vary across zip codes distinguished by, say, per-capita income, but less ideal for identifying the aggregate consumption affects (across all zip codes) of changing home values. Home values at the aggregate level could be responding to an increase in aggregate consumption as opposed to influencing it.

The shock restricted SVAR methodology offers a potential resolution to this dilemma without requiring the use of an instrument that may not be credibly exogenous. To illustrate this point, we consider a bivariate VAR and use shock-based restrictions to winnow the set of solutions that are consistent with mutually uncorrelated shocks to housing and consumption.

Consider an SVAR with two variables: $\mathbf{X}_t = (\Delta \ln C_t, \Delta \ln H_t)'$, where $\Delta \ln C_t$ is the log change in real personal consumption expenditures for the aggregate U.S. economy, and $\Delta \ln H_t$ is the

31

log change in real housing wealth (deflated by PCE deflator). Our objective is to identify a set of mutually uncorrelated structural shocks

$$\mathbf{e}_t = (e_{C,t}, e_{H,t})',$$

where $e_{C,t}$ refers to the consumption shock, and $e_{H,t}$ to the housing shock. Movements in $e_{H,t}$ represent unforecastable changes in housing wealth that are not a response to changes in consumption. These shocks may therefore be used to estimate the dynamic causal effects of an impulse to housing wealth on aggregate consumption.

The possible role of housing wealth in driving aggregate fluctuations has become a subject of special interest as economists ponder the possible factors that lead to the Great Recession, a protracted economic downturn that overlapped with a global financial crisis and a dramatic boom-bust cycle in residential real estate prices. It is now widely accepted that the period of rapid home price appreciation from 2002 to 2006 was associated with a wide-spread relaxation of mortgage lending standards accompanied by declining credit spreads, lower financial market risk premia, and easier financial conditions. Conversely, the housing bust that started some time in 2007 accompanied by a global financial crisis was associated with a sharp increase in credit spreads and a tightening of financial conditions.[11]

With this historical backdrop in mind, we consider a shock-restricted identification strategy to recover $\mathbf{e}_t = (e_{C,t}, e_{H,t})'$. For this application we use a single external variable inequality constraint that requires housing shocks to have a non-zero correlation of the "right" sign with measures of credit spreads or financial conditions. We use two separate external variables for this purpose, employed as alterative measures of financial conditions one at a time. The first, denoted EBP_t is the EBP variable used in the previous application. The second, denoted FCI_t, is the National Financial Conditions Index constructed by the Chicago Federal Reserve. The first is a measure of risk premia in credit markets; the second is a measure of the tightness of financial conditions.

5.1   Shock-Based Constraints

According to the economic history just described, house price movements should be correlated with aggregate financial conditions, with positive (negative) house price shocks associated with looser (tighter) financial conditions and narrower (wider) credit spreads. This motivates the following restrictions.

We consider several sets of alternative restrictions. The first row refers to the covariance restrictions alone. The restrictions in the rows labeled Recursive (a) and Recursive (b) refer

---

[11]For empirical evidence on lending standards during this period see (Favilukis, Kohn, Ludvigson and Van Nieuwerburgh (2013) and Jordà, Schularick and Taylor (2017)). For a theoretical model that shows how relaxed lending standards lower risk premia and raise house prices, see (Favilukis, Ludvigson and Van Nieuwerburgh (2017)).

Table 3: Identification Restrictions, Housing Application

Model	Restrictions	Notes
$\bar{g}_Z = 0$	$\mathrm{vec}(\hat{\Omega}) - \mathrm{vech}(BB') = 0$	covariance restrictions
Recursive (a)	Recursive Identification	variable order: $\mathbf{X}_t = (\Delta \ln C_t, \Delta \ln H_t)'$
Recursive (b)	Recursive Identification	variable order: $\mathbf{X}_t = (\Delta \ln H_t, \Delta \ln C_t)'$
SEE(a)	$\bar{g}_Z = 0$	covariance restrictions
	$\bar{g}_C(e_{H,t}; S_t)$: $\mathrm{corr}(\mathrm{EBP}_t, e_{H,t}) < 0$	external variable: Excess Bond Premium
SEE(b)	$\bar{g}_Z = 0$	covariance restrictions
	$\bar{g}_C(e_{H,t}; S_t)$: $\mathrm{corr}(\mathrm{FCI}_t, e_{H,t}) < 0$	external variable: Financial Condition Index
ADRR	$\bar{g}_Z = 0$	covariance restrictions
	$\sum_{\tau \in \bar{\tau}} e_{H,\tau} < 0$	sign restriction during $\bar{\tau} = [2007\text{:}Q3, 2007\text{:}Q4]$
	$|\mathcal{H}_{C,H,\bar{\tau}}| > |\mathcal{H}_{C,C,\bar{\tau}}|$	Type A/B restriction on historical decomposition

Notes: The table summarizes restrictions in different constrained solution sets. ADRR's Type A and B restrictions on historical decomposition are identical under bivariate VAR system. $\mathcal{H}_{C,H,\bar{\tau}}$ is the contribution of e_H to the unexpected change in log consumption during $\bar{\tau} = [2007\text{:}Q3, 2007\text{:}Q4]$. $q_\alpha(x)$ refers to the α percentile value of x.

to the two possible recursive identification schemes that may be employed for this system. In SEE(a) and SEE(b) we apply the external variable restriction using *either* EBP_t or FCI_t. That is, we apply the constraints one at a time to investigate the sensitivity of the findings to a particular S_t. The restriction requires that a *positive* housing wealth shock be associated with *lower* credit market risk premia and *looser* financial conditions. Conversely, a negative housing wealth shock must be associated with higher risk premia and tighter financial conditions. Note that there is nothing in the external variable constraints that explicitly precludes a non-zero correlation between the consumption shock $e_{C,t}$ and these measures of financial conditions. Nor is there anything that restricts the relative importance of the two shocks in any particular episode.

Finally, the row labeled ADRR considers the "narrative sign" type restrictions introduced in Antolín-Díaz and Rubio-Ramírez (2018). As discussed above, the restrictions considered in Antolín-Díaz and Rubio-Ramírez (2018) differ from those here in two ways. First, they entertain event restrictions that play up the role of some shocks while simultaneously playing down the role of others (e.g., their "Type A and B" restrictions on the historical decompositions), similar in spirit to the traditional narrative-external instrument approach. Second, they do not use external variables at all. Thus to use their approach we need at least one event. For this purpose we consider the U.S. banking panic that occurred in the third and fourth quarters of 2007. The historical chronology of this panic and its specific link to an adverse housing "shock" has been extensively documented by Gorton (2008), Gorton (2009), and Gorton and Metrick (2012), who argue that the panic began August 2007 and continued for the rest of that year. The constraints labeled ADRR impose both a sign restriction on $e_{H,t}$ in 2007:Q3 and

2007:Q4 requiring that the sum of the housing shocks in these two periods be negative, which is to say below average, as well as a Type A/B historical decomposition restriction, requiring that contribution of $e_{H,t}$ to the observed log change in consumption in 2007:Q3 and 2007:Q4 exceed that of $e_{C,t}$.[12] Note that this latter event constraint differs from those considered in the previous applications since it explicitly requires some shocks to play smaller roles than others in this episode.

5.2   Results for the Housing Application

For this application our data are quarterly and we use the largest common sample available for observations on consumption growth, housing wealth growth, and the financial variables used as external variables. This leaves us with a sample that spans the period 1973:Q1 to 2016:Q4. A more detailed description of the data and our sources is provided in the Online Appendix. We estimate the VAR using four quarterly lags. Since the data for this VAR are in log differences, the IRF figures below display the *cumulative* responses of consumption growth to a one standard deviation increase in $e_{H,t}$.

Figure 9 shows the solution sets of IRFs. We focus first on the sets of responses obtained by imposing *only* the covariance structure restrictions $\bar{g}_Z(\mathbf{B}) = 0$. In the absence of additional identifying assumptions it is difficult to assign an interpretation to the shocks, but the case is useful as a benchmark.

Next we add external variable restrictions to shrink the admissible set. The IRFs obtained by imposing SEE(a) are shown in the top panel; those obtained by imposing SEE(b) are shown in the bottom panel. We see that the restrictions greatly narrowed the range of possible responses compared to those based only on $\bar{g}_Z(\mathbf{B}) = 0$, and in both cases they show that a positive housing wealth shock drives up consumption sharply and persistently, with all solutions in the identified sets displaying this pattern. The bounds using EBP_t are particularly tight. The results imply that when a positive housing wealth shock is associated with declining credit spreads and looser financial conditions, or conversely when a negative housing wealth shock is associated with rising credit spreads and tighter financial conditions, it affects aggregate consumer spending. This aspect of the findings echo those in Krishnamurthy and Muir (2017). The application shows that, even without valid instruments, fairly clear conclusions can be drawn about the causal effects of housing on aggregate consumption.

The IRFs obtained under the ADRR restrictions are presented in red dashed lines of Figure 9. The sets are much wider than those under the SEE restrictions. The upper bound of the ADRR solution set is close to those obtained under either the SEE(a) or SEE(b) solution sets, but the lower bound is much lower than those using SEE(a) and SEE(b). This example

---

[12]Due to the use of a bivariate VAR, their Type A and B restrictions are identical for this application.

Figure 9: Housing Wealth and Consumption
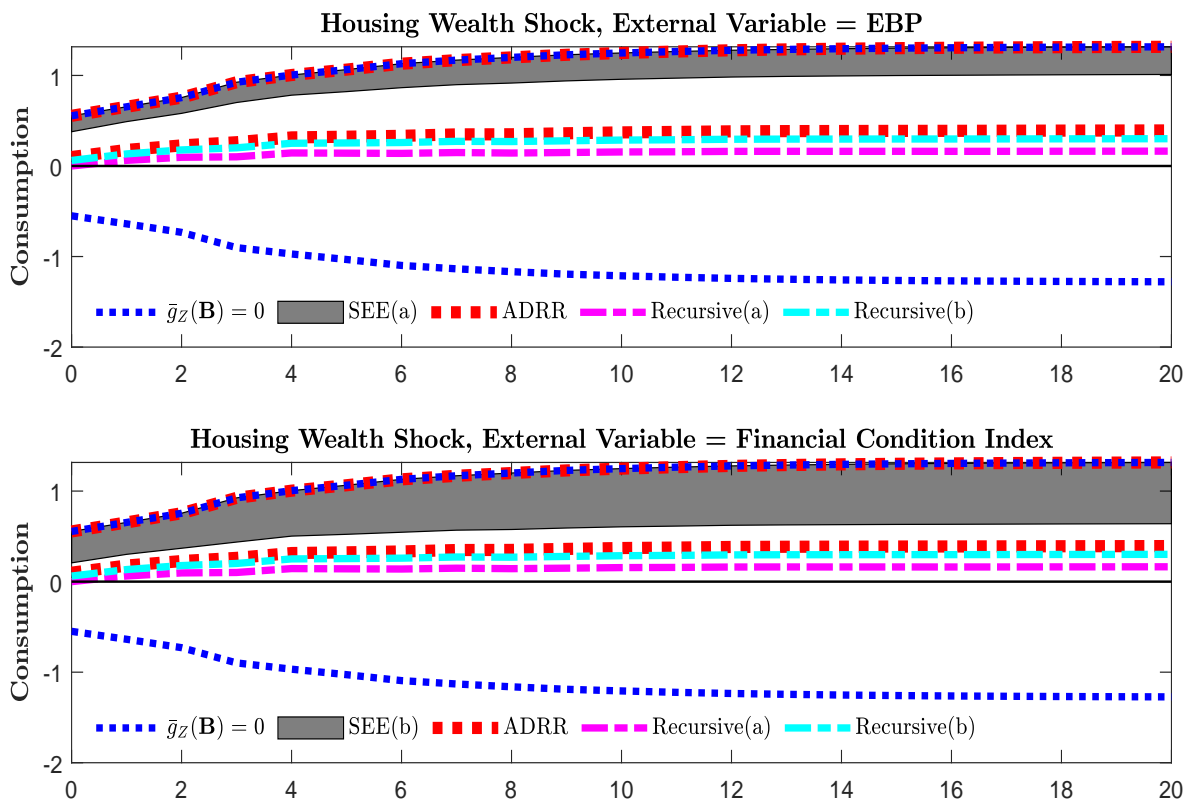
The figure reports the solution set of impulse response to positive, one standard deviation change in $e_{H,t}$ for system $\mathbf{X}_t = (\Delta \ln C, \Delta \ln H)'$ under SEE(a) and SEE(b) restrictions listed in Table 3. The pink dotted line corresponds to the recursive identification with order $\mathbf{X}_t = (\Delta \ln C, \Delta \ln H)'$. The cyan dotted line corresponds to the recursive identification with order $\mathbf{X}_t = (\Delta \ln H, \Delta \ln C)'$. The red dotted line corresponds to the set of solutions under ADRR restriction. The sample spans the period 1973:Q5 to 2016:Q4.

underscores the value for some applications of using information in external variables to help shrink the set of plausible solutions. For this particular application, a single external variable constraint requiring merely that housing shocks exhibit a negative correlation tight financial conditions evidently has more identifying power than a pair of event constraints that restrict the relative importance of housing and consumption shocks in the banking panic of 2007.

It is instructive to also consider the IRFs that would be obtained if this system were point-identified using a recursive identification scheme. Note that the assumptions embedded in the shock-restricted SVAR do not rule out the possibility of a recursive structure, so the estimation is free to recover one if such a structure is consistent with the restrictions. With two variables in the SVAR, there are two possible recursive orderings, hence two possible point estimates of **B**. The IRFs implied by these two point-identified models are shown in Figure 9 as different colored dashed lines.

The results show that, under either recursive ordering, the impact of a housing shock on con-

sumption is estimated to be close to zero at most horizons, implying that housing shocks have little role to play in aggregate consumption fluctuations. Furthermore, both point-identified solutions are outside the range of values suggested by the shock-restricted SVAR. An examination shows why: under either recursive ordering, the identified $e_{H,t}$ is found to be *positively* rather than negatively correlated with both EBP_t and FCI_t, a property that is ruled out by assumption by the SEE restrictions. Given our ex post understanding of events in a sample that included unusually large house price fluctuations, it is doubtful that a positive correlation between these variables is sensible. To have confidence in those identification schemes, one would have to believe that positive housing wealth shocks that were explicitly not a response to changes in consumption coincided with higher risk premia and tighter financial conditions, while negative housing wealth shocks coincided with a relaxation of lending standards and looser financial conditions.

We close this section by considering whether there are important asymmetries in the effects of housing shocks on aggregate consumption over time across the housing boom and bust. For the purposes of this investigation, we define the housing boom period to be 2002:Q1 to 2006:Q4, and the bust period to be 2007:Q1-2011:Q4. Our analysis makes use of an historical decomposition, which measures the contribution of a shock to the observed unexpected change in a variable in the VAR at any given time period t of the sample, or between any two time periods t and $t+h$. For each solution k in the solution set, we calculate the historical contribution of $e_{H,t}$ to the observed unexpected change in $\ln C_t$ at time t and then locate the smallest and largest contributions among all solutions. Figure 10 reports the largest and smallest contributions of both $e_{H,t}$ and $e_{C,t}$ to the cumulative unexpected change in $\ln C_t$ over the two subsamples: 2002:Q1-2006:Q4 and 2002:Q1-2011:Q4. For both plots, the results pertain to the solution set obtained by imposing the SEE(a) restrictions or the ADRR restrictions. Details on how the contributions are computed are given in the Online Appendix.

Consider first the SEE(a) restrictions. As Figure 10 indicates, housing shocks played a substantial role in driving aggregate consumption fluctuations in the housing bust, but their role in the housing boom is unclear. For the boom period, the demeaned log difference in consumption was just 0.20% at an annual rate, while the contribution of $e_{H,t}$ ranges anywhere from -0.27% and 0.28%. The role of the consumption shocks in driving consumption during the boom is also unclear, with the contribution of $e_{C,t}$ ranging from -0.08% to 0.47%. This may be because consumption growth was fairly typical in this subperiod, exceeding its historical mean by only a small magnitude. By contrast, for the period 2007:Q1-2011:Q4, the demeaned log difference in consumption was -2.22% at an annual rate with the estimated contribution of $e_{H,t}$ ranging from -2.32% to -1.31%, indicating that housing shocks explained anywhere from 105% to 60% of the aggregate consumption decline during the housing bust. The range of contributions of the consumption shocks, on the other hand, ranges from -0.91% to 0.10%, making their role

Figure 10: Historical Decomposition of Consumption
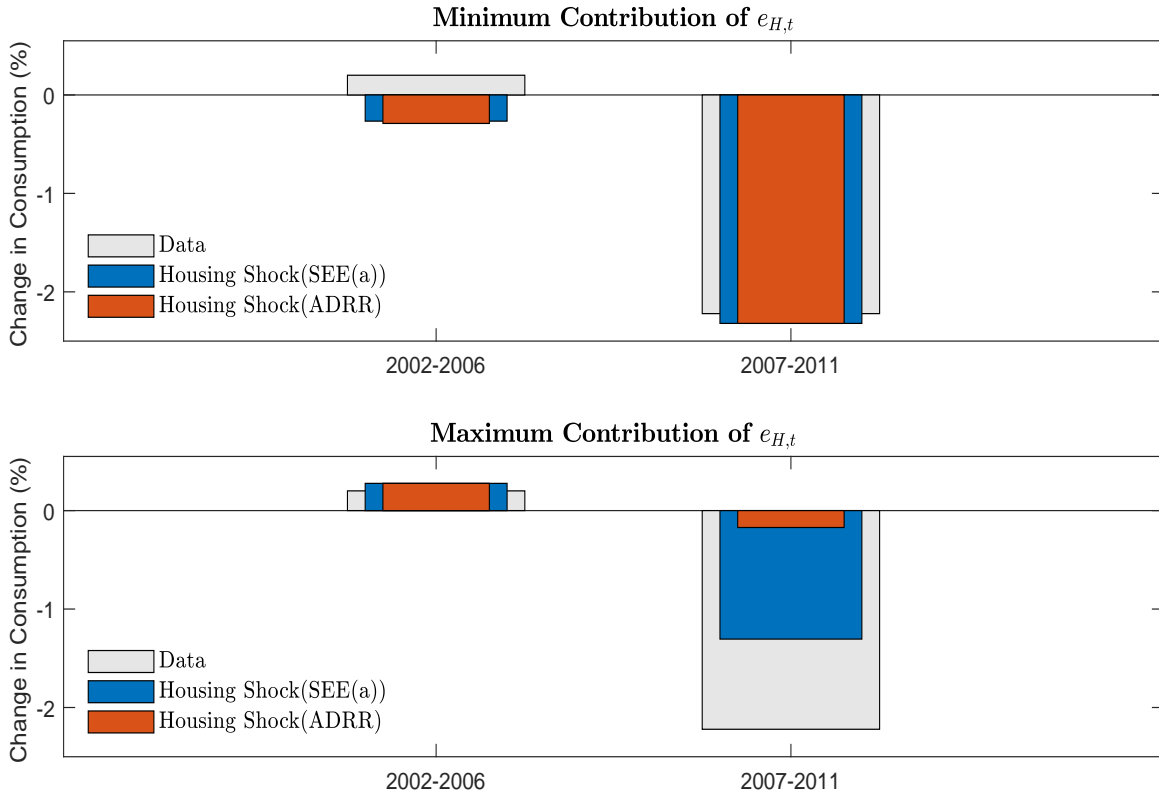
The figure reports the contribution of the housing shock to the changes in demeaned real log consumption from 2002:Q1 to 2006:Q4 and from 2007:Q1 to 2011:Q4 for system $\mathbf{X}_t = (\Delta \ln C, \Delta \ln H)'$ under SEE(a) and ADDR restrictions listed in Table 3. The blue (red) bars in the top panel report the smallest contributions of $e_{H,t}$ to the cumulative unexpected change in $\ln C_t$ over the two subsamples under SEE(a) (ADRR) restriction. The bottom panel reports the largest such contributions. The grey bars in both panels correspond to the actual changes in demeaned log consumption, reported at an annual rate. The VAR parameters and historical decompositions are estimated on a sample that spans the period 1973:Q1 to 2016:Q4.

again unclear. For the entire boom-bust subsample 2002:Q1 to 2011:Q4 where variation was dominated by the bust, the demeaned log difference in consumption was -1.01% at an annual rate and the contribution of housing shocks ranges from -1.12% to -0.79%. We conclude that housing played a potentially large role in the sharp declines in aggregate consumption that were characteristic of the Great Recession, but an unclear role in the housing boom years.

The results using the ADRR restrictions are less clear overall. The findings are similar to those under the SEE(a) restrictions for the boom period. In this case the contribution of $e_{H,t}$ ranges anywhere from -0.29% and 0.28%, while the contribution of $e_{C,t}$ ranges from -0.08% to 0.49%. But in contrast to the results under the SEE(a) restrictions, the ADRR restrictions suggest that a wide range of contributions from housing shocks were possible during the bust period, with the contribution of $e_{H,t}$ ranging from -2.32% to -0.17%, indicating that housing shocks explained anywhere from 8% to 105% of aggregate consumption fluctuations. This

37

contrast with the results using the SEE restrictions is not surprising given the wider solution sets shown in Figure 9.

6   Conclusion

Identifying assumptions need to be imposed in order to give impulse responses generated by vector autoregressions an economically meaningful interpretation. But in many cases the assumptions required for point identification may be overly strong, while commonly used set-identified approaches are often not rich enough to conclude that the data are consistent with a clear causal pattern among the variables. In this paper we explore the properties of a new type of restrictions based on moment inequalities that can help winnow the range of plausible solutions. The restrictions take the form of constraining the structural shocks rather than the parameters of a VAR. To illustrate the potential uses of the approach, we consider two types of restrictions adapted to three applications. The first type restricts the sign and magnitude of identified shocks in specific episodes of history so that they accord with a broad historical understanding of events at particular points in the sample. These restrictions are tantamount to creating dummy variables from the timing of specific events and then putting restrictions on their correlation with the identified shocks, thereby turning a restriction based on a small number of observations into a moment restriction. The second type restricts the correlations between the identified shocks and variables that are external to the VAR. Monte Carlo methods are used to get a sense of sampling variability. The issue of how to best conduct frequentist inference in shock-restricted SVARs remains an important topic for future research.

Most existing empirical work using vector autoregressions focuses closely on the set of theoretical restrictions that may be defensibly placed on the SVAR parameters to achieve identification. But the structural shocks implied by the chosen identification scheme, which are the product of both the data and the restricted parameter vector, are rarely scrutinized. We contend that this focus often misses valuable clues from the data that may be highly informative about the shocks of interest, for two reasons. First, the applications studied here demonstrate that the shocks are additional empirical objects of interest, no matter what the identification scheme. In some cases, the shocks obtained by identification schemes derived exclusively from attention to parameter restrictions appear implausible given an ex post understanding of major historical or economic events. Second, any model that may be written as a VAR provides $n(n+1)/2$ restrictions from the covariance structure, thereby yielding a set of admissible shocks that are a product of the data alone. This set can be employed directly as part of the shock-restricted approach discussed here to provide additional evidence on major economic events of the sample, before any identifying restrictions are imposed.

38

References

Andrews, Donald W. K. and Gustavo Soares, "Inference for parameters defined by moment inequalities using generalized moment selection," *Econometrica*, 2010, *78* (1), 119–157.

Antolín-Díaz, Juan and Juan F Rubio-Ramírez, "Narrative sign restrictions for SVARs," *American Economic Review*, 2018, *108* (10), 2802–29.

Baumeister, Christiane and James D Hamilton, "Structural Interpretation of Vector Autoregressions with Incomplete Identification: Setting the Record Straight," *manuscript, UC San Diego*, 2019.

—— **and Lutz Kilian**, "Forty years of oil price fluctuations: Why the price of oil may still surprise us," *The Journal of Economic Perspectives*, 2016, *30* (1), 139–160.

Cieslak, Anna and Andreas Schrimpf, "Non-monetary news in central bank communication," *Journal of International Economics*, 2019, *118*, 293–315.

—— **and Hao Pang**, "Common shocks in stocks and bonds," *Available at SSRN 3409340*, 2019.

Conley, Timothy G, Christian B. Hansen, and Peter E. Rossi, "Plausibly Exogenous," *The Review of Economics and Statistics*, 2012, *94* (1), 260–272.

Favilukis, Jack, David Kohn, Sydney C. Ludvigson, and Stijn Van Nieuwerburgh, "International Capital Flows and House Prices: Theory and Evidence," in Edward L. Glaeser and Todd Sinai, eds., *Housing and the Financial Crisis*, Chicago, IL: University of Chicago Press, 2013, pp. 235–299.

—— **, Sydney C. Ludvigson, and Stijn Van Nieuwerburgh**, "The Macroeconomic Effects of Housing Wealth, Housing Finance and Limited Risk Sharing in General Equilibrium," *Journal of Political Economy*, 2017, *125* (1), 140–223.

Gafarov, Bulat, Matthias Meier, and José-Luis Montiel-Olea, "Projection Inference for Set-Identified SVARs," 2015. Unpublished paper, New York University.

Gagnon, Joseph, Matthew Raskin, Julie Remache, and Brian Sack, "The financial market effects of the Federal Reserve's large-scale asset purchases," *International Journal of Central Banking*, 2011, *7* (1), 3–43.

Gertler, Mark and Peter Karadi, "Monetary policy surprises, credit costs, and economic activity," *American Economic Journal: Macroeconomics*, 2015, *7* (1), 44–76.

Gilchrist, Simon and Egon Zakrajšek, "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, 2012, *102* (4), 1692–1720.

Gorton, Gary and Andrew Metrick, "Securitized banking and the run on repo," *Journal of Financial economics*, 2012, *104* (3), 425–451.

Gorton, Gary B, "The Panic of 2007," *NBER Working Paper*, 2008, (w14358).

—— , "Slapped in the Face by the Invisible Hand: Banking and the Panic of 2007," *Available at SSRN 1401882*, 2009.

Granziera, Eleonora, Hyungsik Roger Moon, and Frank Schorfheide, "Inference for VARs identified with sign restrictions," *Quantitative Economics*, 2018, *9* (3), 1087–1121.

Hamilton, James D., "Historical Oil Shocks,," in Randall E. Parker and Robert Whaples, eds., *Routledge Handbook of Major Events in Economic History*, New York: Routledge Taylor and Francis Group, 2013, pp. 49–56.

Jordà, Òscar, Moritz Schularick, and Alan M Taylor, "Macrofinancial history and the new business cycle facts," *NBER macroeconomics annual*, 2017, *31* (1), 213–263.

Kilian, Lutz, "Exogenous oil supply shocks: how big are they and how much do they matter for the US economy?," *The Review of Economics and Statistics*, 2008, *90* (2), 216–240.

—— , "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," *American Economic Review*, 2009, *99* (3), 1053–1069.

—— , "Facts and Fiction in Oil Market Modeling," *manuscript, Federal Reserve Bank of Dallas*, 2019.

—— **and Daniel P. Murphy**, "Why agnostic sign restrictions are not enough: understanding the dynamics of oil market VAR models," *Journal of the European Economic Association*, 2012, *10* (5), 1166–1188.

—— **and Daniel P Murphy**, "The role of inventories and speculative trading in the global market for crude oil," *Journal of Applied Econometrics*, 2014, *29* (3), 454–478.

—— **and Helmut Lutkepohl**, *Structural Vector Autoregressive Analysis*, Cambridge University Press, 2016.

Krishnamurthy, Arvind and Annette Vissing-Jorgensen, "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy," *Brookings Papers on Economic Activity*, 2011, *2011* (2), 215–287.

___ **and** ___ , "The ins and outs of LSAPs," in "Kansas City Federal Reserve Symposium on Global Dimensions of Unconventional Monetary Policy" 2013, pp. 57–111.

___ **and Tyler Muir**, "How credit cycles across a financial crisis," Technical Report, National Bureau of Economic Research 2017.

Ludvigson, Sydney C., Sai Ma, and Serena Ng, "Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response?," *American Economic Journal: Macroeconomics,* forthcoming, 2019. http://www.econ.nyu.edu/user/ludvigsons/ucc.pdf.

Mertens, Karel and Morten O Ravn, "The dynamic effects of personal and corporate income tax changes in the United States," *The American Economic Review*, 2013, *103* (4), 1212–1247.

Mian, Atif and Amir Sufi, "House price gains and US household spending from 2002 to 2006," Technical Report, National Bureau of Economic Research 2014.

Montiel-Olea, José-Luis, James H. Stock, and Mark D. Watson, "Inference in Structural VARs with External Instruments," 2015. Unpubished manuscript, Harvard University.

Ramey, Valerie A, "Macroeconomic shocks and their propagation," in "Handbook of Macroeconomics," Vol. 2, Elsevier, 2016, pp. 71–162.

Romer, Christina D and David H Romer, "A new measure of monetary shocks: Derivation and implications," *American Economic Review*, 2004, *94* (4), 1055–1084.

Sims, Christopher A., "Macroeconomics and Reality," *Econometrica*, 1980, pp. 1–48.

Stock, James H. and Mark W. Watson, "NBER Summer Institute Minicourse 2008: What's New in Econometrics: Time Series, Lecture 7: Structural VARs," *Cambridge, Mass.: National Institute for Economic Research. www. nber. org/minicourse_ 2008. html*, 2008.

___ **, Jonathan H. Wright, and Motohiro Yogo**, "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics*, October 2002, *20* (4), 518–528.

Swanson, Eric T, "Measuring the effects of Federal Reserve forward guidance and asset purchases on financial markets," Technical Report, National Bureau of Economic Research 2017.

Uhlig, Harald, "What are the effects of monetary policy on output? Results from an agnostic identification procedure," *Journal of Monetary Economics*, 2005, *52* (2), 381–419.

Zeev, Nadav Ben, "What can we learn about news shocks from the late 1990s and early 2000s boom-bust period?," *Journal of Economic Dynamics and Control*, 2018, *87*, 94–105.

Appendix for Online Publication

Data Description

Oil Application

Data for the system $\mathbf{X}_t = (\Delta prod_t, \, rea_t, \, rpo_t)'$ are obtained from the data file of Kilian (2009) at URL:https://www.aeaweb.org/aer/data/june09/20070211_data.zip. $\Delta prod_t$ is the percentage change in global crude oil production, rea_t is the global demand of industrial commodities variable constructed in Kilian (2009) and rpo_t is the real oil price.

For this application we also use Killian's (2008) measure of "exogenous oil supply shocks" as an external variable. The variable measures shortfalls in OPEC oil production associated with wars and civil disruptions. This indicator is used as an external *instrument* for point identifying oil price shocks in Montiel-Olea et al. (2015).

We use the same data used in Kilian (2009) and the largest common sample period across Kilian (2008) and Kilian (2009) (1973:02-2004:09) for the analysis.

Monetary Policy Application

Both the data for the system $\mathbf{X}_t = \left(i_t^{(12)}, cpi_t, ip_t, ebp_t \right)'$ and the data for the external variable Fed funds futures FF4 are obtained from the data and code file of Gertler and Karadi (2015) at URL:https://www.aeaweb.org/aej/mac/data/0701/2013-0329_data.zip. For the analysis in the paper, we use the same data used by GK and otherwise stick to their sample dates. The sample is monthly and spans the period 1979:07 to 2012:06.

The Greenbook residuals are constructed in Romer and Romer (2004). For each FOMC meeting m, they estimated the following equation,

$$
\begin{aligned}
\Delta TR_m \;=\;& \beta_0 + \beta_1 TR_{m-1} + \sum_{k=-1}^{2} \gamma_k^y g_{m,k} + \sum_{k=-1}^{2} \gamma_k^\pi \pi_{m,k} + \gamma_k^u u_{m,0} \\
&+ \sum_{k=-1}^{2} \lambda_k^y \underbrace{\left(g_{m,k} - g_{m-1,k} \right)}_{\text{Change in growth forecast}} + \sum_{k=-1}^{2} \lambda_k \underbrace{\left(\pi_{m,k} - \pi_{m-1,k} \right)}_{\text{Change in inflation forecast}} + \underbrace{\varepsilon_m}_{\text{Residual}}
\end{aligned}
$$

where ΔTR_m is the change in the target rate at meeting m from the last meeting, $g_{m,k}$ is the k-horizon forecast of real output growth released at meeting m, $\pi_{m,k}$ is the k-horizon forecast of inflation released at meeting m, and $u_{m,0}$ is the current forecast of the unemployment rate at meeting m.

We obtained the Greenbook historical forecast data on real GDP, inflation, and unemployment rate from the Philadelphia Fed's Greenbook Data Set (URL:https://www.philadelphiafed.org/research-and-data/real-time-center/greenbook-data/philadelphia-data-set). We replicated the Romer and Romer (2004) Residuals over the sample 1979:07 to 1996:12. We then used the same estimation equation and extended the residuals to 2012:06.

Housing Application

The total PCE data are from the U.S. Bureau of Economic Analysis (Series ID DPCERX1). The housing wealth combines two series from the Flow of Funds. One is the "Households; owner-occupied real estate" (Flow of Funds B.101 Line 4). The second one is the "Nonfinancial noncorporate business; residential real estate at market value" (Flow of Funds B.104 Line 4). The housing wealth is deflated using the PCE deflator. The excess bonds premium data are obtained from URL: `http://people.bu.edu/sgilchri/Data/GZ_M_August_2016.csv.zip`. The National Financial Conditional Index (FCI) series constructed by the Federal Reserve Bank of Chicago are obtained from FRED (Series ID NFCI) available at URL: `https://fred.stlouisfed.org/series/NFCI`. By construction, higher values of the FCI are associated with tighter financial conditions.

Repeated Sampling Simulation

This appendix describes a bootstrap Monte Carlo procedure to assess the sampling error of our inequality restrictions when $\mathbf{S}_t$ are variables external to the SVAR.

Let R be the number of replications in a repeated sampling experiment. Let "hats" denote estimated values from historical data, e.g., $\widehat{\mathbf{e}}_t$ denotes estimated structural shocks and $\widehat{\mathbf{B}}$ estimated structural covariance matrix. To denote simulated data, we use a "*", while to denote estimated values from simulated data, a "hat" is combined with a "*".

To generate samples of the structural shocks from this solution in a way that ensures the events that appear in historical data also occur in our simulated samples, we first divide the historical sample into two subsamples, one consisting of observations on dates for which event inequality constraints are imposed consisting of observations on non-event dates. For latter subsample consisting of non-event dates, we draw randomly with replacement from the sample estimates of the shocks in the non-event subsample, i.e., from $\{\widehat{\mathbf{e}}_t\}$ $\forall t \notin \bar{\boldsymbol{\tau}} = (\bar{\tau}_1, ..., \bar{\tau}_E)'$. For the former subsample, we divide the set of shocks on each event date $\bar{\tau}_i$ into two categories, those that are unconstrained, denoted $\widehat{\mathbf{e}}_{U,\bar{\tau}_i}$, and those that are subject to the event inequality constraint, denoted $\widehat{\mathbf{e}}_{C,\bar{\tau}_i}$. For $\widehat{\mathbf{e}}_{U,\bar{\tau}_i}$, we draw randomly with replacement from the set of corresponding shocks in the subsample of non-event dates, i.e., $\{\widehat{\mathbf{e}}_{U,t}\}$ $\forall t \notin \bar{\boldsymbol{\tau}} = (\bar{\tau}_1, ..., \bar{\tau}_E)'$, where $\bar{\boldsymbol{\tau}}$ is the vector of event dates and E denotes the number of events. For $\widehat{\mathbf{e}}_{C,\bar{\tau}_i}$, we draw randomly with replacement from the set of corresponding shocks in the subsample of event dates, i.e., from the $\widehat{\mathbf{e}}_{C,\bar{\tau}_i} \in \hat{\mathcal{B}}$ that satisfy the event inequality constraint given $\bar{k}_i$.

Since we identify a set of estimated parameters $\widehat{\mathbf{B}}$ and therefore a set of estimated shocks $\widehat{\mathbf{e}}_t$, we generate R samples of data from each $\widehat{\mathbf{e}}_t$ in the set. This is then repeated for every solution/shock sequence in the solution set to obtain a confidence region for the solution set of impulse responses.

Let M be the number of solutions in the solution set $\bar{\mathcal{B}}(\mathbf{B}; \bar{\mathbf{k}}, \bar{\boldsymbol{\tau}}, \bar{\lambda}, \mathbf{S})$ and let m index an arbitrary solution in the set. Index each draw from the estimated shocks with r and denote the r-th draw from the m-th solution as $\mathbf{e}_t^{mr}$. Each $\mathbf{e}_t^{mr}$ is combined with the $\mathbf{B}$ parameters of the m-th solution, $\widehat{\mathbf{B}}^m$ to generate R samples of size T of $\boldsymbol{\eta}_t^{mr*} = \widehat{\mathbf{B}}^m \mathbf{e}_t^{mr}$. Next, R new samples of $\mathbf{X}_t$ are recursively generated for each replication $r = 1, ..., R$ using $\mathbf{X}_t = \sum_{j=1}^{p} \widehat{\mathbf{A}}_j \mathbf{X}_{t-j} + \boldsymbol{\eta}_t^{mr*}$, with initial conditions fixed at their sample values, $[\mathbf{X}_{-p+1}, ..., \mathbf{X}_0]$. Using each of these new samples of $\mathbf{X}_t$, we fit a VAR(p) model to obtain new least squares estimates $\left[\widehat{\boldsymbol{\eta}}_t^{mr*}, \widehat{\mathbf{A}}_1^{mr*}, ..., \widehat{\mathbf{A}}_p^{mr*}\right]$ and $\widehat{\boldsymbol{\Omega}}^{mr*} = \text{cov}(\widehat{\boldsymbol{\eta}}_t^{mr*}, \widehat{\boldsymbol{\eta}}_t^{mr*})$, and $\hat{\mathcal{B}}^{mr*} = \{\widehat{\mathbf{B}}^{mr*} = \widehat{\mathbf{P}}^{mr*}\mathbf{Q} : \mathbf{Q} \in \mathbb{O}_n, \text{diag}\left(\widehat{\mathbf{B}}^{mr*}\right) \geq 0, \bar{g}_Z(\mathbf{B}) = \mathbf{0}\}$, where where $\mathbb{O}_n$ is the set of $n \times n$ orthonormal matrices and $\widehat{\mathbf{P}}^{mr*}$ is the unique lower triangular Cholesky factor of $\widehat{\boldsymbol{\Omega}}^{mr*}$.

To generate samples of the external variable S_t from m-th solution in a way that ensures that the correlations with the uncertainty shocks that appear in our historical data also appear in our simulated samples, we first generate historical idiosyncratic oil shortfall shocks e_{St}^m as the fitted residuals from regressions of S_t on a single autoregressive lag and on $\widehat{\mathbf{e}}_t$, respectively. Next, we draw randomly with replacement from e_{St}^m with the exception that, as above, we fix the values for these shocks in each replication in the periods $\bar{\tau}$, to obtain $r = 1, ..., R$ new values e_{St}^{mr} and R new values of S_t by recursively iterating on

$$S_t^{mr} = d_{01}^m + \widehat{\rho} S_{t-1}^{mr} + \mathbf{d}^{m\prime} \mathbf{e}_t^{mr} + e_{St}^{mr} \tag{5}$$

with initial condition fixed at its initial sample values, S_1. The parameters $\widehat{\rho}$ are estimated from a first order autoregression for S_t. The parameters $\mathbf{d}^{m\prime}$ in (5) are calibrated to target the observed correlations $\text{corr}(S_t, \widehat{\mathbf{e}}_t^m)$ for the m-th solution in historical data so that $\text{corr}(S_t^{mr}, \mathbf{e}_t^{mr})$ roughly equal $\text{corr}(S_t, \widehat{\mathbf{e}}_t^m)$ on average across all replications R.

We construct confidence sets for the set of IRFs in repeated samples as follows. The number of replications is set to $R = 1,000$. In each replication of each solution, $K = 1.5$ million rotation matrices $\mathbf{Q}$ are entertained, but only $K_{mr} \leq K$ rotations will generate solutions that are admitted into the solution set for that replication, $\bar{\mathcal{B}}^{mr*}(\cdot)$. Let $\Theta_{i,j,s}^{m,r,k}$ be the s-period ahead response of the i-th variable to a standard deviation change in shock j at the k-th rotation of K_{mr}, for replication r and solution m.[13] Let $\underline{\Theta}_{i,j,s}^{m,r} = \min_{k \in [1, K_{mr}]} \Theta_{i,j,s}^{m,r,k}$ and $\overline{\Theta}_{i,j,s}^{m,r} = \max_{k \in [1, K_{mr}]} \Theta_{i,j,s}^{m,r,k}$. Each $(\underline{\Theta}_{i,j,s}^{m,r}, \overline{\Theta}_{i,j,s}^{m,r})$ pair represents the extreme (highest and lowest) dynamic responses in replication r of solution m. From the quantiles of the set $\left\{\underline{\Theta}_{i,j,s}^{m,r}\right\}_{m=1, r=1}^{M, R}$ that includes all replications

---

[13]The s-period ahead dynamic responses to one-standard deviation shocks in the jth variable are defined as

$$\frac{\partial \mathbf{X}_{t+s}}{\partial e_{jt}} = \widehat{\boldsymbol{\Psi}}_s^{mr*} \widehat{\mathbf{b}}^{mrkj*},$$

where $\widehat{\mathbf{b}}^{mrkj*}$ is the jth column of $\widehat{\mathbf{B}}^{mrk*}$ and the coefficient matrixes $\widehat{\boldsymbol{\Psi}}_s^{mr*}$ are given by $\widehat{\boldsymbol{\Psi}}^{mr*}(L) = \widehat{\boldsymbol{\Psi}}_0^{mr*} + \widehat{\boldsymbol{\Psi}}_1^{mr*} L + \widehat{\boldsymbol{\Psi}}_2^{mr*} L^2 + \ldots = \widehat{\mathbf{A}}^{mr*}(L)^{-1}$.

for all solutions we can obtain the $\alpha/2$ critical point $\underline{\Theta}_{i,j,s}(\alpha/2)$. Similarly, from the quantiles of $\{\overline{\Theta}_{i,j,s}^{m,r}\}_{m=1,r=1}^{M,R}$, we have the $1 - \alpha/2$ critical point $\overline{\Theta}_{i,j,s}(1 - \alpha/2)$. Eliminating the lowest and highest $\alpha/2$ percent of the samples gives a $(1 - \alpha)\%$ percentile-based confidence interval defined by

$$CI_{\alpha,g} = \left[\underline{\Theta}_{i,j,s}(\alpha/2), \ \overline{\Theta}_{i,j,s}(1 - \alpha/2)\right].$$

$CI_{\alpha,g}$ denotes the confidence intervals for sets of solutions that satisfy all constraints, including the event and correlation constraints: $\bar{g}_Z(\mathbf{B}) = 0$, $\bar{g}_E(\mathbf{B}; \bar{\tau}, \bar{\mathbf{k}}) \geq 0$, $\bar{g}_C(\mathbf{B}; \mathbf{S}) \geq 0$. We use CI_{α,g_Z} to denote the confidence intervals for sets of solutions that satisfy only the covariance structure restrictions $\bar{g}_Z(\mathbf{B}) = 0$.

Historical Decomposition

The decompositions are computed as follows. Let $\Theta_{i,j,s}^k$ be the s-period ahead impulse response of the i-th variable to a standard deviation change in shock j at the k-th solution of the solution set. For each solution k in the solution set, we calculate the historical contribution $\mathcal{H}_{C,H,t}^k$ of $e_{H,t}$ to the observed unexpected change in log consumption at time t as

$$\mathcal{H}_{C,H,t}^k = \sum_{s=0}^{t-1} \Theta_{C,H,s}^k e_{H,t-s}^k.$$

Let $\underline{\mathcal{H}}_{C,H,t} = \min_{k\in[1,K]} \mathcal{H}_{C,H,t}^k$ and $\overline{\mathcal{H}}_{C,H,t} = \max_{k\in[1,K]} \mathcal{H}_{C,H,t}^k$. The $(\underline{\mathcal{H}}_{C,H,t}, \overline{\mathcal{H}}_{C,H,t})$ pair gives the smallest and largest contributions, respectively, of $e_{H,t}$ to the unexpected change in log consumption among all solutions in the solution set. The contributions for the subsamples 2002:Q1-2006:Q4 and 2002:Q1-2011:Q4 are computed by cumulating the time t contributions over these subsamples. Note that since the shocks are mean zero by construction, the data must be demeaned in order to make the contributions meaningful.