

# Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response?

SYDNEY C. LUDVIGSON\*

NYU and NBER

SAI MA<sup>†</sup>

NYU

SERENA NG<sup>‡</sup>

Columbia University and NBER

First draft: March 2015

This draft: January 4, 2017

## Abstract

Uncertainty about the future rises in recessions. But is uncertainty a source of business cycle fluctuations or an endogenous response to them, and does the type of uncertainty matter? We find that sharply higher uncertainty about real economic activity in recessions is most often an endogenous response to other shocks that cause business cycle fluctuations, while uncertainty about financial markets is a likely source of the fluctuations. To establish the dynamic effects of uncertainty shocks, we exploit information from external variables and the timing of extraordinary economic events to identify structural vector autoregressions under credible interpretations of the structural shocks.

JEL: G11, G12, E44, E21.

Keywords: uncertainty, volatility, economic fluctuations, business cycles

---

\*Department of Economics, NYU, 19 W. 4th St, 6th Floor, New York, NY 10012. (sydney.ludvigson@nyu.edu)

<sup>†</sup>Department of Economics, NYU, 19 W. 4th St, 6th Floor, New York, NY 10012 (sai.ma@nyu.edu)

<sup>‡</sup>Department of Economics, 420 West 118th St., New York, NY 10027. (serena.ng@columbia.edu)

Ludvigson acknowledges financial support from the C.V. Starr Center for Applied Economics at NYU. Ng acknowledges support from the National Science Foundation under grant SES-0962431 and SES-1558623. We thank seminar participants at the 2016 ME, EFFE, and EF&G group meetings at the NBER, the CEPR-DRG-SAFE Conference on Banking, Monetary Policy, and Macroeconomic Performance in Frankfurt May 2016, the New Developments in Business Cycle Analysis conference in Rome June 2016, the 2016 Econometric Society meetings, the Board of Governors, Collegio Carlo Alberto, NYU, and UC Davis, and Yale for helpful comments. We are grateful to Ian Dew-Becker, José Luis Montiel Olea, Mikkel Plagborg-Møller, Jim Stock, and Stephen Terry for many valuable discussions.

# 1 Introduction

A large literature in macroeconomics investigates the relationship between uncertainty and business cycle fluctuations. Interest in this topic has been spurred by a growing body of evidence that uncertainty rises sharply in recessions. This evidence is robust to the use of specific proxy variables such as stock market volatility and forecast dispersion as in Bloom (2009), or a broad-based measure of macroeconomic uncertainty, as in Jurado, Ludvigson, and Ng (2015) (JLN hereafter). But while this evidence substantiates a role for uncertainty in deep recessions, the question of whether uncertainty is an exogenous source of business cycle fluctuations or an endogenous response to economic fundamentals is not fully understood. Existing results are based on convenient but restrictive identifying assumptions and have no explicit role for financial markets, even though the uncertainty measures are strongly correlated with financial market variables. This paper considers a novel identification strategy to disentangle the causes and consequences of real and financial uncertainty.

The question of causality and the identification of exogenous variation in uncertainty is a long-standing challenge of the uncertainty literature. The challenge arises in part because there is no theoretical consensus on whether the uncertainty that accompanies deep recessions is primarily a cause or effect (or both) of declines in economic activity. Theories in which uncertainty is defined as the time varying volatility of a fundamental shock cannot address this question because, by design, there is no feedback response of uncertainty to other shocks if the volatility process is specified to evolve exogenously. And, obviously, models in which there is no exogenous variation in uncertainty cannot be used to analyze the direct effects of uncertainty shocks. It is therefore not surprising that many theories for which uncertainty plays a role in recessions reach contradictory conclusions on this question, as we survey below.

A separate challenge of the uncertainty literature pertains to the origins of uncertainty. Classic theories assert that uncertainty originates from economic fundamentals such as productivity, and that such real economic uncertainty, when interacted with market frictions, discourages real activity. But some researchers have argued that uncertainty dampens the economy through its influence on financial markets (e.g., Gilchrist, Sim, and Zakrajsek (2010)). Moreover, as surveyed by Ng and Wright (2013), all the post-1982 recessions have origins in financial markets, and these recessions have markedly different features from recessions where financial markets play a passive role. From this perspective, if financial shocks are subject to time-varying volatility, financial market uncertainty—as distinct from real economic uncertainty—could be a key player in recessions, both as a cause and as a propagating mechanism. The Great Recession of 2008, characterized by sharp swings in financial markets, hints at such a linkage. Yet so far the literature has not disentangled the contributions of real versus financial uncertainty to business cycle fluctuations.

Econometric analyses aimed at understanding the role of uncertainty for business cycle fluctuations face their own challenges, especially when the body of theoretical work does not provide precise identifying restrictions for empirical work. Attempts to identify the “effects” of uncertainty shocks in existing empirical work are primarily based on recursive schemes within the framework of vector-autoregressions (VAR).<sup>1</sup> But studies differ according to whether uncertainty is ordered ahead of or after real activity variables in the VAR. While a recursive structure is a reasonable starting point, any presumed ordering of the variables is hard to defend on theoretical grounds given the range of models in the literature. Contemporaneous changes in uncertainty can arise both as a cause of business cycle fluctuations and as a response to other shocks. Recursive structures explicitly rule out this possibility since they presume that some variables respond only with a lag to others.

It is with these challenges in mind that we return to the questions posed above: is uncertainty primarily a source of business cycle fluctuations or a consequence of them? And what is the relation of real versus financial uncertainty to business cycle fluctuations? The objective of this paper is to establish a set of stylized facts that addresses these questions econometrically, against which a wide range of individual models could be evaluated. To do so, we employ a small-scale structural vector autoregression (SVAR). To confront the challenges just discussed, we take a two-pronged approach. First, our empirical analysis explicitly distinguishes *macro* uncertainty from *financial* uncertainty. The baseline SVAR we study describes the dynamic relationship between three variables: an index of macro uncertainty,  $U_{Mt}$ , a measure of real economic activity,  $Y_t$  (e.g., production, employment), and a new financial uncertainty index introduced here,  $U_{Ft}$ . Second, rather than relying on ordering assumptions for identification, we use a different identification scheme that is less restrictive, both because it allows for simultaneous feedback between uncertainty and real activity, and because it can be used to test whether a lower recursive structure is supported by the data.

In conventional SVAR analyses, the focus is typically on the identified impulse response functions and decomposition of variances. The properties of the shocks are often not scrutinized even though the stated purpose of a SVAR analysis is to identify the structural shocks. In our analysis, the shocks play a central role. Our identification strategy is to complement standard covariance restrictions with implicit or explicit restrictions on the shocks. We impose two types of shock-based restrictions. The first requires that the identified shocks be consistent with economic reasoning in a small number of extraordinary events, such as the 1987 stock market crash and the financial crisis/Great Recession of 2007-09. We refer to these as “event timing constraints,” or simply *event constraints*.

The second type of constraints is referred to as *component correlation constraints*, so called

---

<sup>1</sup>See Bachmann, Elstner, and Sims (2013), Bloom (2009), Bloom (2014), Bekaert, Hoerova, and Duca (2013), Gilchrist, Sim, and Zakrajsek (2010), and JLN.

because restrictions are placed on the correlation between the identified shocks and certain sub-components of special variables denoted  $S_t$ . These  $S_t$  variables are external to the SVAR system and are not themselves valid instruments. But each variable in  $S_t$  is presumed to have a component that is correlated with both types of uncertainty shocks yet contemporaneously uncorrelated with real activity shocks, while having another component that is correlated with financial uncertainty shocks but contemporaneously uncorrelated with both real activity and macro uncertainty shocks. Since these components behave as if they were valid instruments, we refer to them as “synthetic proxies.” Restricting the correlation between these synthetic proxies and the shocks is akin to requiring an instrument to be minimally relevant.

The event and component correlation constraints are used to constrict the number of solutions in an identified set that can otherwise be arbitrarily large. Naturally, we may have many solutions or no solution depending on the implementation. As we shall see, our proposed constraints are capable of substantially winnowing the set of identified impulse responses. A key part of the analysis is to find observables external to our SVAR that are driven by a multitude of innovations, including the two types of uncertainty shocks that we are interested in. We argue below that both theory and evidence suggest that aggregate stock market returns are natural candidates for such  $S_t$  variables. Our maintained economic hypothesis is that stock market returns should be correlated with both types of uncertainty shocks and therefore have the two components described above.

The empirical exercise additionally requires that appropriate measures of macro and financial uncertainty be available. To this end, we exploit a data rich environment, working with 134 macro monthly time series and 147 financial variables. The construction of macro uncertainty follows JLN. The same approach is used to construct a broad-based measure of financial uncertainty that has never been used in the literature. Macro uncertainty is itself an aggregate of uncertainties in variables from three categories: real activity, price, and financial. To better understand the contributions of each of these categories, we also replace  $U_{Mt}$  in the VAR with an uncertainty measure based on the real activity sub-component. Uncertainty about real activity is of special interest because classic uncertainty theories postulate that uncertainty shocks have their origins in economic fundamentals and hence should show up as uncertainty about real economic activity.

Before summarizing our main results, it should be made clear that the structural shocks we identify do not necessarily correspond to primitive shocks of any particular model, as this is not our goal. Our real activity shocks could originate from technology, monetary policy, preferences, or government expenditure innovations, and our uncertainty shocks could originate from economic policies and/or technology. Our approach explicitly eschews imposing a specific model structure. It is instead designed to elicit the dynamic causal relationships between business cycle and uncertainty fluctuations when commonly used timing, ordering, or other

restrictions valid only under special theoretical assumptions are difficult to defend.

Our main results may be stated as follows. First, positive shocks to financial uncertainty are found to cause a sharp decline in real activity that persists for many months, lending support to the hypothesis that heightened uncertainty is an exogenous impulse that causes recessions. These effects are especially large for several measures of real activity, notably production, employment and a broad real activity index. The finding that heightened uncertainty has negative consequences for real activity is qualitatively similar to that of preexisting empirical work that uses recursive identification schemes (e.g., Bloom (2009), JLN), but differs in that we trace the source of this result specifically to broad-based financial market uncertainty rather than to various uncertainty proxies or broad-based macro uncertainty. We also show that the converse is not supported by our evidence: exogenous shocks to real activity have no clear effect on financial uncertainty given the set of SVAR parameters we identify.

Second, the identification scheme used here reveals something new that is not possible to uncover under recursive schemes: macro and financial uncertainty have a very different dynamic relationship with real activity. Specifically, unlike financial uncertainty, sharply higher macro and real activity uncertainty in recessions is found to be an endogenous response to business cycle fluctuations. That is, negative economic activity shocks are found to cause increases in both macro uncertainty and in the sub-index that measures uncertainty about real economic activity, but there is no evidence that independent shocks to macro or real activity uncertainty cause lower economic activity. Indeed the opposite is often true: exogenous shocks to both macro and real uncertainty are found to *increase* real activity, consistent with “growth options” theories discussed below.

Third, our results are distinct from those obtained using recursive identification. Under any recursive ordering of the variables in our VAR, exogenous shocks that increase macro or real uncertainty appear to reduce real activity, in a manner that is qualitatively similar to financial uncertainty shocks. This result does not hold in the less restrictive SVAR studied here and appears to be an artifact of invalid timing assumptions under recursive identification. Further investigation reveals that the SVAR we study reflects a non-zero contemporaneous correlation between  $U_{Ft}$  and  $Y_t$ , as well as between  $U_{Mt}$  and  $Y_t$ , which is inconsistent with any recursive ordering. Tests of the validity of a recursive structure are easily rejected by the data.

The rest of this paper is organized as follows. Section 2 reviews related literature and provides motivation for our maintained economic hypothesis that stock market returns have components that are correlated with macro and financial uncertainty shocks but contemporaneously uncorrelated with real activity, and also correlated with financial uncertainty but contemporaneously uncorrelated with both real activity and macro uncertainty. Section 3 details the econometric framework and identification employed in our study, describes how the synthetic proxies are constructed, and discusses the data and empirical implementation. In this

section we also show how, with some additional restrictions, our approach can be interpreted as the output of a system estimation for a larger VAR that includes both  $\mathbf{X}_t = (U_{Mt}, Y_t, U_{Ft})$ , and  $S_t$ . Section 4 presents empirical results using broad based macro uncertainty  $U_{Mt}$ , while Section 5 reports results for systems that isolate the sub-component of  $U_{Mt}$  corresponding to real activity variables. Section 6 reports results for additional cases including alternative bounds, an assessment of the validity of recursive identification restrictions, and system estimation results for a VAR in  $(\mathbf{X}_t, S_t)$ . Section 7 summarizes and concludes. A large number of additional results are reported in an online Appendix. The underpinnings of the identification scheme as well as Monte Carlo simulations calibrated to the present application are further explored in Ludvigson, Ma, and Ng (2016).

## 2 Related Literature

A large literature addresses the question of uncertainty and its relation to economic activity.<sup>2</sup> Besides the evidence cited above for the U.S., Nakamura, Sergeyev, and Steinsson (2012) estimate growth rate and volatility shocks for 16 developed countries and find that they are substantially negatively correlated. Theories for which uncertainty plays a key role differ widely on the question of whether this correlation implies that uncertainty is primarily a cause or a consequence of declines in economic activity. In most cases, it is modeled either as a cause or an consequence, but not both.

The first strand of the literature proposes uncertainty as a cause of lower economic growth. This includes models of the real options effects of uncertainty (Bernanke (1983), McDonald and Siegel (1986)), models in which uncertainty influences financing constraints (Gilchrist, Sim, and Zakrajsek (2010), Arellano, Bai, and Kehoe (2011)), or precautionary saving (Basu and Bundick (2012), Leduc and Liu (2012), Fernández-Villaverde, Pablo Guerrón-Quintana, and Uribe (2011)). These theories almost always presume that uncertainty is an exogenous shock to some economic fundamental. Some theories presume that higher uncertainty originates directly in the process governing technological innovation, which subsequently causes a decline in real activity (e.g., Bloom (2009), Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2012)).

A second strand of the literature postulates that higher uncertainty arises solely as a *response* to lower economic growth, emphasizing a variety of mechanisms. Some of these theories suggest that bad times incentivize risky behavior (Bachmann and Moscarini (2011), Fostel and Geanakoplos (2012)), or reduce information and with it the forecastability of future outcomes (Van Nieuwerburgh and Veldkamp (2006) Fajgelbaum, Schaal, and Taschereau-Dumouchel (2014)), or provoke new and unfamiliar economic policies whose effects are highly uncertain (Pástor and Veronesi (2013)), or create a greater misallocation of capital across sectors (Ai, Li,

---

<sup>2</sup>This literature has become voluminous. See Bloom (2014) for a recent review of the literature.

and Yang (2015)), or generate endogenous countercyclical uncertainty in consumption growth because investment is costly to reverse (Gomes and Schmid (2016)).

And yet a third literature has raised the possibility that some forms of uncertainty can actually *increase* economic activity. “Growth options” theories of uncertainty postulate that a mean-preserving spread in risk generated from an unbounded upside coupled with a limited downside can cause firms to invest and hire, since the increase in mean-preserving risk increases expected profits. Such theories were often used to explain the dot-com boom. Examples include Bar-Ilan and Strange (1996), Pastor and Veronesi (2006), Kraft, Schwartz, and Weiss (2013), Segal, Shaliastovich, and Yaron (2015).

This brief review reveals a rich literature with a wide range of predictions about the relationship between uncertainty and real economic activity. Yet the absence of a theoretical consensus on this matter, along with the sheer number of theories and limited body of evidence on the structural elements of specific models, underscores the extent to which the question of cause and effect is fundamentally an empirical matter that must be settled in an econometric framework with as little specific theoretical structure as possible, so that the various theoretical possibilities can be nested in empirical tests. Commonly used recursive identification schemes cannot achieve this objective, since by construction they rule out the possibility that uncertainty and real activity could influence one another within the period. Our econometric model nests any recursive identification scheme, so we can test whether such timing assumptions are plausible. We find they are rejected by the data.

Our maintained hypothesis that stock market returns should be correlated with uncertainty shocks builds on work in asset pricing emphasizing the idea that stock market variation is the result of several distinct (and orthogonal) sources of stochastic variation. For example, one quantitatively important component is attributable to *acyclical* risk premia variation, and more generally appears to be uncorrelated with most measures of real activity. This component is valuable for our objective because it is exogenous to real activity, but may still be relevant for both macro and financial uncertainty. Yet another component could be attributable to fluctuations in factors like corporate leverage, the risk-bearing capacity of financial intermediaries, or the risk aversion or “sentiment” of market participants that may be correlated with the volatility of the stock market. In equilibrium asset pricing models, if leverage increases or the risk-bearing capacity of intermediaries declines, volatility of the corporate sector’s equity return increases. Thus changes in factors like leverage, intermediary risk-bearing capacity (and possibly changes in risk aversion or sentiment) are likely to be correlated with financial uncertainty, but may have little to do with uncertainty about economic fundamentals. This component is valuable for our objective because it is plausibly uncorrelated with both real activity and uncertainty about economic fundamentals, but may still be relevant for financial market uncer-

tainty.<sup>3</sup> Consistent with the existence of this type of component, JLN document that there are many spikes in stock market uncertainty that do not coincide with an important movement in either real activity or macro uncertainty. These findings motivate our maintained hypothesis that measures of equity market returns are promising non-uncertainty variables comprised of several distinct sources of stochastic variation, two of which are useful for identifying the shocks of interest.

Our use of external variables  $S_t$  is related to a recent line of econometric research in SVARs that uses information contained in external instruments to identify structural dynamic causal effects.<sup>4</sup> Of these, Stock and Watson (2012) study uncertainty shocks, using a measure of stock market volatility and/or a news media measure of policy uncertainty from Baker, Bloom, and Davis (2013), as separate external instruments for identifying the effects of uncertainty shocks in a SVAR. Our study differs in some fundamental ways. First, our approach relies on a set of economic assumptions that is distinct from that of standard IV approach, hence the moment conditions used to identify the model parameters and shocks are not the same. The identification strategy in Stock and Watson (2012) for uncertainty shocks presumes that the external variables themselves (i.e., stock market volatility, policy uncertainty) are valid instruments, correlated with the uncertainty shock of interest but not with the other shocks. By contrast, our approach explicitly views both the stock market and our uncertainty measures as partly endogenous, forcing us to confront the identification quandary. Our identification assumption is instead that aggregate stock market returns contain components that are correlated with the structural uncertainty shocks and we consider an identification strategy that relies on lower bounds for the absolute values of these correlations. Second, Stock and Watson (2012) focus exclusively on identifying the effects of uncertainty shocks and do not attempt to identify all shocks in the system, including the effects of real activity on uncertainty.

Berger, Dew-Becker, and Giglio (2016) take a different approach. Using options data they find that bad times are associated with higher realized volatility but not higher expected volatility, a result that they interpret as consistent with the hypothesis that higher uncertainty is a consequence of negative economic shocks rather than a cause. This interpretation is not intended to provide an explicit identification of uncertainty shocks, however.

Finally, Baker and Bloom (2013), who use disaster-like events as instruments for stock market volatility with the aim of isolating exogenous variation in uncertainty. This has some

---

<sup>3</sup>For empirical evidence that risk premia in financial markets have an important component uncorrelated with real activity, see Bianchi, Ilut, and Schneider (2014); Lettau and Ludvigson (2013), Greenwald, Lettau, and Ludvigson (2014), Kozak and Santosh (2014), and Muir (2014). Theoretical examples in which real or financial uncertainty varies independently of economic fundamentals include Bansal and Yaron (2004); Wachter (2013); Gourio (2012); Bianchi, Ilut, and Schneider (2014); Gabaix and Maggiori (2013); He and Krishnamurthy (2013).

<sup>4</sup>See for example Hamilton (2003), Kilian (2008), Mertens and Ravn (2013); Stock and Watson (2008), Stock and Watson (2012), and Olea, Stock, and Watson (2015).



similarities with our approach, in that it implicitly assumes that certain components of stock market fluctuations (those associated with “disasters”) are exogenous. Whereas disasters chosen subjectively are presumed to be valid instruments for uncertainty, we instead use external stock return data and unusual events to constrain a set of estimable moment restrictions. It is of interest that we arrive at complementary conclusions, despite the differing methodologies for identifying exogenous variation.

### 3 Econometric Framework

This section outlines our econometric approach. Subsection 1 explains the identification strategy. Subsections 2 and 3 explain estimation methodology and the uncertainty measures. Subsection 4 shows how our approach can be interpreted as the output of a restricted system estimation for a larger VAR that includes both  $\mathbf{X}_t$  and  $S_t$ .

#### 3.1 The SVAR and Identification

Let  $\mathbf{X}_t$  denote a  $K \times 1$  vector time series. We suppose that  $\mathbf{X}_t$  has a reduced-form vector autoregressive and an infinite-order moving average representation given respectively by:

$$\mathbf{X}_t = \mathbf{k} + \mathbf{A}_1 \mathbf{X}_{t-1} + \mathbf{A}_2 \mathbf{X}_{t-2} + \cdots + \mathbf{A}_p \mathbf{X}_{t-p} + \boldsymbol{\eta}_t. \quad (1)$$

$$\mathbf{X}_t = \boldsymbol{\mu} + \boldsymbol{\Psi}(L) \boldsymbol{\eta}_t \quad (2)$$

$$\boldsymbol{\eta}_t \sim (0, \boldsymbol{\Omega}), \quad \boldsymbol{\Omega} = \mathbb{E}(\boldsymbol{\eta}_t \boldsymbol{\eta}_t')$$

where  $\boldsymbol{\Psi}(L) = \mathbf{I}_n + \boldsymbol{\Psi}_1 L + \boldsymbol{\Psi}_2 L^2 + \dots$  is a polynomial in the lag operator  $L$  of infinite order, and  $\boldsymbol{\Psi}_s$  is the  $(n \times n)$  matrix of coefficients for the  $s$ th lag of  $\boldsymbol{\Psi}(L)$ . The reduced form innovations  $\boldsymbol{\eta}_t$  are related to the structural shocks  $\mathbf{e}_t$  by an invertible  $K \times K$  matrix  $\mathbf{H}$ :

$$\boldsymbol{\eta}_t = \mathbf{H} \boldsymbol{\Sigma} \mathbf{e}_t \equiv \mathbf{B} \mathbf{e}_t \quad (3)$$

$$\mathbf{e}_t \sim (0, \mathbf{I}_K), \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & 0 & \cdot & 0 \\ 0 & \sigma_{22} & 0 & 0 \\ 0 & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \sigma_{KK} \end{pmatrix}, \quad \sigma_{jj} \geq 0 \quad \forall j. \quad (4)$$

where  $\mathbf{B} \equiv \mathbf{H} \boldsymbol{\Sigma}$ . The structural shocks  $\mathbf{e}_t$  are mean zero with unit variance, and are serially and mutually uncorrelated. A normalization is required to pin down the sign and scale of the shocks. We adopt the unit effect normalization

$$\text{diag}(\mathbf{H}) = 1. \quad (5)$$

The objective of the exercise is to study the dynamic effects and the relative importance of each structural shock  $j$ . These are summarized by the impulse response function (IRF)  $\frac{\partial \mathbf{X}_{t+s}}{\partial e_{jt}} = \boldsymbol{\Psi}_s \mathbf{b}^j$

(where  $\mathbf{b}^j$  is the  $j$ th column of  $\mathbf{B}$ ) and the fraction of  $s$ -step ahead forecast error variance of  $\mathbf{X}_t$  that is attributable to each structural shock. The SVAR identification problem concerns identifying the elements of  $\mathbf{H}$  and  $\mathbf{\Sigma}$ , from which the structural IRFs and variance decompositions are computed.

To study the impulse and propagating mechanism of uncertainty shocks while explicitly distinguishing between macro and financial market uncertainty, we consider a system with  $K = 3$  variables. Our baseline SVAR is based on  $\mathbf{X}_t = (U_{Mt}, Y_t, U_{Ft})'$ , where  $U_{Mt}$  denotes macro uncertainty,  $Y_t$  denotes a measure of real activity, and  $U_{Ft}$  denotes financial uncertainty. The corresponding reduced form shocks  $\boldsymbol{\eta}_t = (\eta_{Mt}, \eta_{Yt}, \eta_{Ft})'$  are related to the three structural form shocks  $\mathbf{e}_t = (e_{Mt}, e_{Yt}, e_{Ft})'$  for macro uncertainty, real activity, and financial uncertainty, as follows:

$$\begin{aligned}\eta_{Mt} &= B_{MM}e_{Mt} + B_{MY}e_{Yt} + B_{MF}e_{Ft} \\ \eta_{Yt} &= B_{YM}e_{Mt} + B_{YY}e_{Yt} + B_{YF}e_{Ft} \\ \eta_{Ft} &= B_{FM}e_{Mt} + B_{FY}e_{Yt} + B_{FF}e_{Ft},\end{aligned}$$

where  $B_{ij}$  is the element of  $\mathbf{B}$  that gives the contemporaneous effect of the  $j$ th structural shock on the  $i$ th variable. The standard covariance restrictions come from the covariance structure of  $\boldsymbol{\eta}_t$ , which provides  $K(K + 1)/2 = 6$  equations in  $\mathbf{B}$ :

$$\text{vech}(\boldsymbol{\Omega}) = \text{vech}(\mathbf{B}\mathbf{B}'), \quad (6)$$

where  $\text{vech}(\boldsymbol{\Omega})$  stacks the unique elements of the symmetric matrix  $\boldsymbol{\Omega}$ . There are nine unknown elements in  $\mathbf{B}$ .

To motivate our procedure, it is helpful to begin by considering the external instrumental variables (IV) approach where valid instruments are observed. To do so, suppose for the moment that we have measures of  $Y_t$ ,  $U_{Mt}$ ,  $U_{Ft}$ , and two valid external instruments  $Z_{1t}$  and  $Z_{2t}$  satisfying the following:

**Assumption A:** Let  $\mathbf{Z}_t = (Z_{1t}, Z_{2t})'$  be two instrumental variables such that

$$\begin{aligned}(A.i) \quad & \mathbb{E}[Z_{1t}e_{Mt}] \neq 0, \quad \mathbb{E}[Z_{1t}e_{Yt}] = 0, \quad \mathbb{E}[Z_{1t}e_{Ft}] \neq 0 \\ (A.ii) \quad & \mathbb{E}[Z_{2t}e_{Mt}] = 0, \quad \mathbb{E}[Z_{2t}e_{Yt}] = 0, \quad \mathbb{E}[Z_{2t}e_{Ft}] \neq 0.\end{aligned}$$

Assumption A are conditions for instrument exogeneity and relevance.  $Z_{1t}$  is an instrument that is correlated with both macro and financial uncertainty, but contemporaneously uncorrelated (exogenous) with respect to real activity.  $Z_{2t}$  is an instrument that is correlated with financial uncertainty, but contemporaneously uncorrelated (exogenous) with respect to macro uncertainty and real activity.

Let  $\mathbf{m}_{1t} = (\text{vech}(\boldsymbol{\eta}_t \boldsymbol{\eta}_t'), \text{vec}(\mathbf{Z}_t \otimes \boldsymbol{\eta}_t))'$  and  $\boldsymbol{\beta} = \text{vec}(\mathbf{B})$ . At the true value of  $\boldsymbol{\beta}$ , denoted  $\boldsymbol{\beta}^0$ , the SVAR external IV model satisfies

$$0 = \mathbb{E}[\mathbf{g}_1(\mathbf{m}_{1t}; \boldsymbol{\beta}^0)], \quad (7)$$

written out in full as follows:

$$\begin{aligned} 0 &= \text{var}(\eta_M) - B_{MM}^2 + B_{MY}^2 + B_{MF}^2 \\ 0 &= \text{var}(\eta_Y) - B_{YM}^2 + B_{YY}^2 + B_{YF}^2 \\ 0 &= \text{var}(\eta_F) - B_{FM}^2 + B_{FY}^2 + B_{FF}^2 \\ 0 &= \text{cov}(\eta_M, \eta_Y) - B_{MM}B_{YM} + B_{MY}B_{YY} + B_{MF}B_{YF} \\ 0 &= \text{cov}(\eta_Y, \eta_F) - B_{YM}B_{FM} + B_{YY}B_{FY} + B_{FF}B_{YF} \\ 0 &= \text{cov}(\eta_M, \eta_F) - B_{MM}B_{FM} + B_{MY}B_{FY} + B_{MF}B_{FF} \\ 0 &= B_{MF}\mathbb{E}[Z_{2t}\eta_Y] - B_{YF}\mathbb{E}[Z_{2t}\eta_{Mt}] \\ 0 &= B_{FF}\mathbb{E}[Z_{2t}\eta_{Yt}] - B_{YF}\mathbb{E}[Z_{2t}\eta_{Ft}] \\ 0 &= (B_{MM}B_{FF} - B_{MF}B_{FM})\mathbb{E}[Z_{1t}\eta_{Yt}] - (B_{YF}B_{FM} - B_{YM}B_{FF})\mathbb{E}[Z_{1t}\eta_{Mt}] \\ &\quad - (B_{MM}B_{YF} - B_{MF}B_{YM})\mathbb{E}[Z_{1t}\eta_{Ft}]. \end{aligned}$$

With two external instruments, the model has nine equations in nine unknowns. The first six are from the covariance structure. The next two equations are due to the three moments implied by Assumption (A.ii). The final equation is due to the three moments implied by Assumption (A.i). It is straightforward to prove that, under Assumption A, the unit-effect normalization, and the restrictions on the admissible parameter space,  $\boldsymbol{\beta}$  is point-identified. (See the online Appendix.) In essence, identification in this IV analysis is achieved by (i) using movements in  $U_{Mt}$  and  $U_{Ft}$  that are correlated with  $Z_{1t}$  to identify the effects of uncertainty shocks and disentangle them from shocks to real activity, (ii) using movements in  $U_{Ft}$  that are correlated with  $Z_{2t}$  to identify the effects of  $U_{Ft}$  shocks and disentangle them from macro uncertainty shocks, and (iii) using movements in  $Y_t$  that are uncorrelated with both  $Z_{1t}$  and  $Z_{2t}$  to identify the effects of real activity shocks and disentangle them from uncertainty shocks.

Since we take the stand in this application that our uncertainty measures are potentially endogenous, it is then natural to ask why we do not simply find observable instruments. One answer is that credible external instruments for uncertainty shocks that are truly exogenous may be difficult or impossible to find and defend. Indeed, existing uncertainty proxies are likely to be among the variables that fall into this category. But even if the exogeneity assumption is questionable, variables external to the VAR could still contain valuable information about the parameters and shocks of interest. The next subsection exploits this idea by proposing a methodology to construct synthetic proxy variables that should contain at least some information for uncertainty shocks, information that, once combined with additional economic

restrictions, can help narrow the set of plausible solutions and the range of dynamic causal effects in the SVAR for  $\mathbf{X}_t$ .

### 3.2 Construction of Synthetic Proxies

Suppose that the ideal instruments  $Z_{1t}$  and  $Z_{2t}$  have no credible observable counterparts. The next step is to develop a methodology to construct synthetic proxies in the spirit of such variables. To motivate our method, recall that two stage least squares uses projections to purge the endogenous variations from a relevant regressor. Our approach is similar except that we purge the endogenous variations from observed variables  $S_t$ . The residuals are our synthetically created proxies for  $Z_{1t}$  and  $Z_{2t}$ . We denote these constructed components  $\mathbf{Z}_t(\boldsymbol{\beta})$  to emphasize that they are functions of parameters  $\boldsymbol{\beta}$  to be estimated.

In the present context, we make use of observable variables  $S_t$  that are driven by the structural shocks  $\mathbf{e}_t = (e_{Yt}, e_{Mt} \text{ and } e_{Ft})'$ , as well as other shocks collected into an  $e_{St}$  that are uncorrelated with  $\mathbf{e}_t$ . A theoretical premise of the paper is that structural uncertainty shocks should be reflected in stock prices. Thus we use measures of stock market returns as  $S_t$ . Under this maintained assumption, we may represent  $S_t$  as

$$S_t = \delta_0 + \delta_Y Y_t + \delta_M U_{Mt} + \delta_F U_{Ft} + \delta_S(L) S_{t-1} + \delta_X(L)' \mathbf{X}_{t-1} + e_{St} \quad (8)$$

where  $\mathbf{X}_t = (Y_t, U_{Mt}, U_{Ft})'$ . The residual  $e_{St}$  could be driven by any number of shocks orthogonal to  $\mathbf{e}_t$ . One interpretation is risk premium shocks driven by factors orthogonal to uncertainty, such as a pure sentiment shock, but the precise interpretation is not important to what follows. It is clear that  $S_t$  and  $\mathbf{X}_t$  are endogenous variables and least squares estimation of (8) will yield inconsistent estimates. However, we are not interested in these parameters. Our interest in stock market returns is solely that they have components that are useful for understanding uncertainty shocks. Consider the projections

$$S_t = d_{10} + d_{1Y} e_{Yt} + d_{1S}(L) S_{t-1} + Z_{1t} \quad (9a)$$

$$S_t = d_{20} + d_{2M} e_{Mt} + d_{2Y} e_{Yt} + d_{2S}(L) S_{t-1} + Z_{2t}, \quad (9b)$$

where  $S_t$  in (9a) and (9b) are not necessarily the same variable. Equation (9a) forms an orthogonal decomposition of  $S_t$  conditional on its own lag into a component that is spanned by  $e_{Yt}$  and a component  $Z_{1t}$  that is orthogonal to  $e_{Yt}$ . Similarly, equation (9b) purges the effect of  $e_{Yt}$  and  $e_{Mt}$  from  $S_t$  conditional on its lags to arrive at  $Z_{2t}$ . Note that  $Z_{1t}$  and  $Z_{2t}$  include the effects of  $\mathbf{X}_{t-1}$ . As  $U_{Mt}$  and  $U_{Ft}$  can be serially correlated, their lagged values can predict future excess stock market returns.

If  $e_Y$  and  $e_M$  were observed, then simple least squares regressions would produce estimates of  $Z_{1t}$  and  $Z_{2t}$  that satisfy Assumption A by construction. But  $\mathbf{e}_t$  themselves depend on  $\boldsymbol{\beta}$ ,

hence  $\mathbf{Z}(\boldsymbol{\beta})$  also depends on  $\boldsymbol{\beta}$ . The problem that this creates is that if  $\hat{\mathbf{Z}}_1'\hat{\mathbf{e}}_Y = 0$ ,  $\hat{\mathbf{Z}}_2'\hat{\mathbf{e}}_Y = 0$ , and  $\hat{\mathbf{Z}}_2'\hat{\mathbf{e}}_M = 0$  for some  $\hat{\boldsymbol{\beta}} = \text{vec}(\hat{\mathbf{B}})$ , any orthonormal rotation of  $\hat{\mathbf{B}}$  to  $\tilde{\mathbf{B}} = \hat{\mathbf{B}}\mathbf{Q}'$  and  $\hat{\mathbf{e}}$  to  $\tilde{\mathbf{e}} = \mathbf{Q}\hat{\mathbf{e}}$  will also result in  $\tilde{\mathbf{Z}}_1'\tilde{\mathbf{e}}_Y = 0$ ,  $\tilde{\mathbf{Z}}_2'\tilde{\mathbf{e}}_Y = 0$ , and  $\tilde{\mathbf{Z}}_2'\tilde{\mathbf{e}}_M = 0$ . In other words, the three exogeneity conditions hold by construction and no longer offer information about  $\boldsymbol{\beta}$ . Hence, the nine moment restrictions in (7) cannot identify the SVAR parameters when the external instruments are replaced by the synthetic proxies. If we collect all the solutions that satisfy (7) into the set  $\hat{\mathcal{B}}$ , this set can be infinitely large.

To address this problem, we abandon the goal of point identification in favor of less restrictive economic assumptions that are supported by the data. Specifically, we consider two types of restrictions that appear new in the SVAR literature. The general idea is that some solutions of  $\boldsymbol{\beta}$  will produce  $\mathbf{Z}(\boldsymbol{\beta})$  and  $\mathbf{e}(\boldsymbol{\beta})$  that are at odds with reasonable economic judgment. We employ two types of *winnowing* constraints to dismiss such solutions in  $\hat{\mathcal{B}}$  to arrive a restricted solution set  $\bar{\mathcal{B}}(\bar{c}, \bar{C}, \bar{k})$ .

**1 Component correlation constraints:** Let  $c_{kj}(\boldsymbol{\beta}) = \text{corr}(Z_{kt}(\boldsymbol{\beta}), e_{jt}(\boldsymbol{\beta}))$  be the sample correlation between  $\mathbf{Z}_k(\boldsymbol{\beta})$  and the shock in  $\mathbf{e}_t(\boldsymbol{\beta}) = (e_{Mt}, e_{Yt}, e_{Ft})$  with label  $j$ .

- i  $|c_{1M}(\boldsymbol{\beta})| > \bar{c}$ ,  $|c_{1F}(\boldsymbol{\beta})| > \bar{c}$ , and  $|c_{2F}(\boldsymbol{\beta})| > \bar{c}$ .
- ii For  $c(\boldsymbol{\beta}) = (c_{1M}(\boldsymbol{\beta}), c_{1F}(\boldsymbol{\beta}), c_{2F}(\boldsymbol{\beta}))'$ ,  $\sqrt{c(\boldsymbol{\beta})'c(\boldsymbol{\beta})} > \bar{C}$ .

**2 Event constraints:** For  $\mathbf{e}_t(\boldsymbol{\beta}) = \mathbf{B}^{-1}\boldsymbol{\eta}_t$  and  $\bar{k} = (\bar{k}_1, \bar{k}_2, \bar{k}_3)'$ ,

- i  $e_{Ft_1}(\boldsymbol{\beta}) > \bar{k}_1$  where  $t_1$  is the period 1987:10 of the stock market crash.
- ii There exists a  $t_2 \in [2007:12, 2009:06]$  such that  $e_{Ft_2}(\boldsymbol{\beta}) > \bar{k}_2$ .
- iii For all  $t_2 \in [2007:12, 2009:06]$ ,  $e_{Yt_2}(\boldsymbol{\beta}) < \bar{k}_3$

The first set of restrictions regards the correlation between uncertainty shocks and aggregate stock market returns. Granted that a non-zero correlation is a maintained assumption, it leaves open the question of how correlated. Rather than taking a stand on a particular magnitude, we set lower bounds on the absolute correlations. This allows us to make use of the fact that the component correlations  $c_{1M} = \mathbf{Z}_1'(\hat{\boldsymbol{\beta}})\hat{\mathbf{e}}_M$ ,  $c_{1F} = \mathbf{Z}_1'(\hat{\boldsymbol{\beta}})\hat{\mathbf{e}}_F$ ,  $c_{2F} = \mathbf{Z}_2'(\hat{\boldsymbol{\beta}})\hat{\mathbf{e}}_F$  are not invariant to orthonormal rotations. That is to say, the component correlations generated by  $\hat{\mathbf{B}}$  will in general be different from those generated by  $\tilde{\mathbf{B}} = \hat{\mathbf{B}}\mathbf{Q}'$ . Hence solutions in the unconstrained set that do not satisfy the lower bound will be dismissed. In implementation, we require that each correlation  $c_{1M}(\hat{\boldsymbol{\beta}})$ ,  $c_{1F}(\hat{\boldsymbol{\beta}})$ ,  $c_{2F}(\hat{\boldsymbol{\beta}})$  individually exceeds a pre-specified  $\bar{c}$ , and collectively exceeds  $\bar{C}$ .

The second set of restrictions requires that the financial uncertainty shocks identified in October 1987 (black Monday) and during the 2007-2009 financial crisis be large and positive,

and that the identified output shocks during the Great Recession not take on large positive values. These event timing restrictions also act to shrink the unconstrained set because, while  $\tilde{e}_t$  and  $\hat{e}_t$  have the same mean and variance,  $\tilde{e}_t \neq \hat{e}_t$  at any particular  $t$ . Solutions that imply favorable financial uncertainty and/or output shocks during these times are dismissed in light of the sub-par economic conditions and/or extreme volatility in the stock market that prevailed. The precise  $t_2$  dates are set in accordance with NBER dating of the Great Recession, which coincides with the timing of the financial crisis.

It remains to discuss the construction of the unconstrained solution set  $\hat{\mathcal{B}}$ . The possible solutions in  $\hat{\mathcal{B}}$  are obtained by initializing  $\mathbf{B}$  to be the lower Cholesky factorization of  $\mathbf{\Omega}$  for an arbitrary ordering of the variables, and then rotating it by  $K = 40,000$  random orthogonal matrices  $\mathbf{Q}$ . Each rotation begins by drawing an  $n \times n$  matrix  $\mathbf{G}$  of NID(0,1) random variables. Then  $\mathbf{Q}$  is taken to be the orthonormal matrix in the  $\mathbf{QR}$  decomposition of  $\mathbf{G} = \mathbf{QR}$  and  $\mathbf{QQ}' = \mathbf{I}_n$ . A  $\beta$  in the unconstrained solution set  $\hat{\mathcal{B}}$  is also in the constrained solution set  $\bar{\mathcal{B}}(\bar{c}, \bar{C}, \bar{k})$  only if the event and component correlation constraints are all satisfied.

Some comments about the implementation of this approach bear discussion. First, the parameter values for the bounds  $\bar{c}$ ,  $\bar{C}$ , and  $\bar{k}$  must in general vary with the data under investigation. It should be clear that if the values for the bounds are overly restrictive, the constrained solution set will be empty, while if they are too unrestrictive the constraints themselves will have no identifying power. Moreover, a particular choice of values for  $\bar{c}$ ,  $\bar{C}$ , and  $\bar{k}$  may be highly restrictive for one system of data but entirely unrestrictive for another. For the applications here, we consider several different systems for  $\mathbf{X}_t$  that vary according to how real activity is measured (production, employment, or a real activity index), or whether macro uncertainty or the sub-index for real activity uncertainty is used. The bounds are adjusted accordingly across these cases, in order for the degree of restrictiveness of the event and correlation constraints to be similar. We discuss the precise values for the bounds below as they pertain to each case.

Second, though no one solution in  $\bar{\mathcal{B}}(\bar{c}, \bar{C}, \bar{k})$  is any more likely than another, it will be useful to have one solution as reference point in the discussions. We use what will be referred to as the ‘max-C’ solution. This is the solution in the restricted set  $\bar{\mathcal{B}}(\bar{c}, \bar{C}, \bar{k})$  with the highest collective correlation  $\sqrt{c(\beta)'c(\beta)}$ , defined as:

$$\hat{\beta}^{\text{max-C}} = \arg \max_{\beta \in \bar{\mathcal{B}}(\bar{c}, \bar{C}, \bar{k})} \sqrt{c(\beta)'c(\beta)}. \quad (10)$$

Third, in conventional IV analyses the instruments are observed and can be readily used to formulate sample orthogonality conditions, as in the nine moment restrictions of (7). If the synthetic variables  $\mathbf{Z}(\beta)$  were to be thought of as synthetic instruments, then conceptually they would need to be constructed prior to estimation. But the construction of  $\mathbf{Z}(\beta)$  itself necessitates values of  $\mathbf{B}$ . In this case, an iterative step would be necessary to ensure that the  $\mathbf{B}$  used to construct  $\mathbf{Z}_t(\beta)$  is consistent with the solution that emerges from (7). This approach was

taken in an earlier draft of the paper, in which the procedure was referred to iterative projection IV. It is now understood that, though conceptually appealing, the additional iteration comes at a computational cost with little to gain, as the results without this step are virtually identical to those with the additional step. It should be noted, however, that regardless of whether this step is taken, the synthetic  $\mathbf{Z}(\boldsymbol{\beta})$  components satisfy the exogeneity restrictions of assumption A, since the structural shocks  $\mathbf{e}_t$  are by construction mutually uncorrelated.

To summarize, identification is predicated on three economic assumptions. First, the external variables  $S_t$  must have components that are relevant for the uncertainty shocks, as specified by the correlation constraints. Second, the identified shocks must be consistent with a priori economic reasoning in a small number of extraordinary events whose interpretation is relatively incontrovertible. Third,  $S_t$  is excluded from the VAR so that its shocks cannot affect the variables in  $\mathbf{X}_t$  either contemporaneously or with a lag. Below we show how this last assumption can be empirically evaluated.

To have confidence in this implementation, Ludvigson, Ma, and Ng (2016) use Monte Carlo experiments to study the properties of the estimator. The results for a data generating process calibrated to the empirical application here shows that the procedure produces solution sets that are substantially narrowed by applying the event and correlation constraints described above.

### 3.3 System Estimation

The estimation procedure just discussed is based on an SVAR for  $\mathbf{X}_t$ . While  $S_t$  plays a role in identification, it is excluded from the SVAR. We refer to the foregoing analysis as the *subsystem approach*. However, it is also possible to apply the event and component correlation constraints to a larger VAR in  $(\mathbf{X}_t, S_t)'$ . We refer to this as the *full system approach*. For this purpose, we consider a single  $S_t$ .

The full system VAR takes the same form as (1); the only difference is that  $S_t$  is now included in the VAR. The reduced form errors for the full system are  $\boldsymbol{\eta}_t = (\eta'_{Xt}, \eta_{St})'$ . The structural shocks are  $(\mathbf{e}'_{Xt} \ e_{St})'$  with  $\boldsymbol{\eta}_t = \mathbf{B}\mathbf{e}_t$ . The  $\mathbf{B}$  matrix now has 16 parameters and the covariance structure gives 10 pieces of information. We assume that the shocks  $e_{St}$  do not contemporaneously affect  $\mathbf{X}_t$ . This means that the impact sub-vector giving the effects of  $e_{St}$  on  $\mathbf{X}_t$ , denoted  $\mathbf{B}_{XS} = (B_{MS}, B_{YS}, B_{FS})'$ , is zero. These three zero restrictions imply

$$\begin{pmatrix} \eta_{Mt} \\ \eta_{Yt} \\ \eta_{Ft} \\ \eta_{St} \end{pmatrix} = \begin{pmatrix} B_{MM} & B_{MY} & B_{MF} & 0 \\ B_{YM} & B_{YY} & B_{YF} & 0 \\ B_{FM} & B_{FY} & B_{FF} & 0 \\ B_{SM} & B_{SY} & B_{SF} & B_{SS} \end{pmatrix} \begin{pmatrix} e_{Mt} \\ e_{Yt} \\ e_{Ft} \\ e_{St} \end{pmatrix}. \quad (11)$$

The synthetic variables  $\mathbf{Z}_t$  are now defined as

$$\begin{aligned} Z_{1t} &= \eta_{St} - B_{SY}e_{Yt} = B_{SM}e_{Mt} + B_{SF}e_{Ft} + B_{SS}e_{St} \\ Z_{2t} &= Z_{1t} - B_{SM}e_{Mt} = B_{SF}e_{Ft} + B_{SS}e_{St}. \end{aligned}$$

Hence they are functions of the structural parameters. This treatment of  $\mathbf{Z}_t$  is conceptually distinct from the subsystem analysis earlier when  $\mathbf{Z}_t$  was treated as a residual from a projection. The full system is estimated using the component correlation and event constraints, just as for the subsystem, except that the residual  $\eta_{St}$  is used to construct  $\mathbf{Z}_t$  in place of  $S_t$ . Hence as in the case with the subsystem analysis, the model is underidentified. To address this problem, we again use the event and component correlation constraints to narrow the set of plausible parameters. In the full system, the correlation constraints are given by

$$\begin{aligned} c_{1M}(\boldsymbol{\beta}) &= \frac{\text{corr}(Z_{1t}, e_{Mt})}{\sigma_{Z_1}} = \frac{B_{SM}}{\sqrt{B_{SM}^2 + B_{SF}^2 + B_{SS}^2}} \\ c_{1F}(\boldsymbol{\beta}) &= \frac{\text{corr}(Z_{1t}, e_{Ft})}{\sigma_{Z_1}} = \frac{B_{SF}}{\sqrt{B_{SM}^2 + B_{SF}^2 + B_{SS}^2}} \\ c_{2F}(\boldsymbol{\beta}) &= \frac{\text{corr}(Z_{2t}, e_{Ft})}{\sigma_{Z_2}} = \frac{B_{SF}}{\sqrt{B_{SF}^2 + B_{SS}^2}}, \end{aligned}$$

where the second equalities follow by recalling that  $e_{Mt}$  and  $e_{Ft}$  have unit standard deviations. Evidently, these correlations explicitly depend on the parameters of the  $S$  equation. Thus, as in the subsystem analysis, they are not invariant to orthonormal rotation of  $\mathbf{e}_X$  and the parameters of the subsystem.

It is of interest to compare the full and subsystem analyses. In the subsystem analysis, the process that generates  $S_t$  is left unspecified. As such, it can be a function of variables other than  $\mathbf{X}_t$ , both contemporaneously, and at lags. By contrast, the full system approach specifies the process for  $S_t$ . Any misspecification in one equation can affect all equations in the system. On the other hand, the full system merely constrains the contemporaneous effect of  $S_t$  on  $\mathbf{X}_t$  to zero. This is weaker than assuming that  $S_t$  is exogenous for  $\mathbf{X}_t$ , which additionally prevents the lags of  $S_t$  from affecting  $\mathbf{X}_t$ . Constraining the current and lagged values of  $S_t$  to zero amounts to the subsystem analysis of excluding  $S_t$  from the larger VAR altogether. It should however be noted that excluding the past values of  $S_t$  from the equations for  $\mathbf{X}_t$  is not needed for the set identification described above. Thus the assumption that  $S_t$  can be excluded from the VAR for  $\mathbf{X}_t$  places overidentifying restrictions on the full system that can be evaluated empirically. A simple way to do so is to compare the impulse response functions estimated for the three variable system  $\mathbf{X}_t = (U_{Mt}, Y_t, U_{Ft})'$  with those from a larger system that includes  $S_t$  but does not restrict the coefficients of  $S_{t-j}$  in the equations for  $\mathbf{X}_t$  to zero, for  $j \geq 1$ . We present these results below.



### 3.4 Measuring Uncertainty and Stock Market Returns

In our estimation we work with several different aggregate measures of uncertainty, which are indexes constructed over individual uncertainties for a large number of observable time-series. A long-standing difficulty with empirical research on this topic has been the measurement of uncertainty. JLN find that common uncertainty proxies contain economically large components of their variability that do not appear to be generated by a movement in genuine uncertainty across the broader economy. This occurs both because these proxies over-weight certain series in the measurement of aggregate uncertainty, and because they erroneously attribute forecastable fluctuations to a movement in uncertainty. Equity market volatility, for example, contains a non-trivial component generated from forecastable variation in stock returns. The estimated macro uncertainty index constructed in JLN is designed to address these issues and improve the measurement of aggregate uncertainty. The methodology used here for constructing uncertainty indexes follows JLN and we refer the reader to that paper for details.

Let  $y_{jt}^C \in Y_t^C = (y_{1t}^C, \dots, y_{N_C t}^C)'$  be a variable in category  $C$ . Its  $h$ -period ahead uncertainty, denoted by  $\mathcal{U}_{jt}^C(h)$ , is defined to be the volatility of the purely unforecastable component of the future value of the series, conditional on all information available. Specifically,

$$\mathcal{U}_{jt}^C(h) \equiv \sqrt{\mathbb{E} \left[ (y_{jt+h}^C - \mathbb{E}[y_{jt+h}^C | I_t])^2 | I_t \right]} \quad (12)$$

where  $I_t$  is information available. If the expectation today of the squared error in forecasting  $y_{jt+h}$  rises, uncertainty in the variable increases. As in JLN, the conditional expectation of squared forecast errors in (12) is computed from a stochastic volatility model, while the conditional expectation  $\mathbb{E}[y_{jt+h}^C | I_t]$  is replaced by a diffusion index forecast, augmented to allow for nonlinearities. These are predictions of an autoregression augmented with a small number of common factors  $q_t = (q_{1t}, \dots, q_{rt})'$  estimated from a large number of economic time series  $x_{it}$  each with factor representation  $x_{it} = \Lambda'_{it} q_t + e_{\chi, it}$ . Nonlinearities are accommodated by including polynomial terms in the factors, and factors estimated squares of the raw data. The use of large datasets reduces the possibility of biases that arise when relevant predictive information is ignored. Let  $Y_t^C = (y_{1t}^C, \dots, y_{N_C t}^C)'$  generically denote the series that we wish to compute uncertainty in.

Uncertainty in category  $C$  is an aggregate of individual uncertainty series in the category :

$$U_{Ct}(h) \equiv \text{plim}_{N_C \rightarrow \infty} \sum_{j=1}^{N_C} \frac{1}{N_C} \mathcal{U}_{jt}^C(h) \equiv \mathbb{E}_C[\mathcal{U}_{jt}^C(h)]. \quad (13)$$

In this paper, we consider four categories of uncertainty:

Category ( $C$ )	$Y_t^C$	$N_C$
(M): Macro	all variables in $\chi^M$	134
(F): Financial	all variables in $\chi^F$	147
(R): Real activity	real activity variables in $\chi^M$	73

We use two datasets covering the sample 1960:07-2015:04.<sup>5</sup> The first is a monthly *macro dataset*,  $\mathcal{X}_t^M$ , consisting of 134 mostly macroeconomic time series take from McCracken and Ng (2016). The second is a *financial dataset*  $\mathcal{X}_t^F$  consisting of a 147 of monthly financial indicators, also used in Ludvigson and Ng (2007) and JLN, but updated to the longer sample. The real uncertainty index  $U_{Rt}$  is an equally-weighted average of the individual uncertainties about 73 series in Groups 1 through 4 of  $\mathcal{X}^M$ . These include output and income variables, labor market measures, housing market indicators, and orders and inventories. Additional predictors for variables in  $\mathcal{X}_{it}^M$  include factors formed from  $\mathcal{X}_{it}^F$  and vice-versa, squares of the first factor of each, and factors in the squares of individual series,  $(\mathcal{X}_{it}^M)^2$  and  $(\mathcal{X}_{it}^F)^2$ .

Our use of stock returns  $S_t$  to generate instruments is grounded in the theoretical premise that both macro and financial uncertainty shocks should be reflected in stock market returns. There is no reason, however, that the regressands in (9a) and (9b) must be exactly the same measure of stock market activity. All measures of stock market activity are highly correlated because they contain a large common component (much of which is orthogonal to the rest of the economy). In order to introduce some additional independent variation in our two instruments, our base cases use different measures of aggregate stock market activity  $S_{1t}$  and  $S_{2t}$ , although in practice we get very similar results if we use the same value-weighted stock market index return in (9a) and (9b). Specifically, for  $S_{1t}$ , the regressand for (9b), we use the Standard and Poor 500 stock market index return. For  $S_{2t}$ , the regressand in (9a), we use  $\alpha_p \text{crsp}_t + (1 - \alpha_p) \text{small}_t$ , which is a portfolio weighted average of the return on the CRSP value-weighted stock index in excess of the one-month Treasury bill rate and the smallest decile stock market return in the NYSE.<sup>6</sup> We set the portfolio weight  $\alpha_p$  to be a value close to one, thereby giving only a small amount of additional weight to small stocks. Small stocks are less representative of the market as a whole, and it is unclear how highly correlated they should be with aggregate uncertainty measures. For the base case results presented below we set  $\alpha_p = 0.94$ . This constructs a portfolio that gives slightly more weight to small stocks than what they receive in the value-weighted CRSP index. However, we also investigated a range of values for  $\alpha_p \in [0.75, 1]$  and found very similar estimates and impulse responses for all weights in this range.

<sup>5</sup> A detailed description of the series is given in the Data Appendix of the online location where updated JLN uncertainty index data are posted: [http://www.sydneyludvigson.com/s/jln\\_data\\_appendix\\_update.pdf](http://www.sydneyludvigson.com/s/jln_data_appendix_update.pdf)

<sup>6</sup> The CRSP index is a value-weighted return of all stocks in NYSE, AMEX, and NASDAQ.

## 4 Results for $\mathbf{X}_t = (U_{Mt}, Y_t, U_{Ft})'$

This section presents empirical results. We begin by studying systems with macro uncertainty  $U_{Mt}$ . We then move on to consider real uncertainty  $U_{Rt}$  formed exclusively from real activity variables.

We consider  $h = 1$  (one-month uncertainty) and several measures of  $Y_t$ . The first two measures are the log of real industrial production, denoted  $ip_t$ , and the log of employment, denoted  $emp_t$ . While industrial production is a widely watched economic indicator of business cycles, it only captures goods-producing industries and has been a declining share of GDP. Employment only covers the labor market. Hence we also consider an additional measure of real activity: the cumulated sum of the first common factor estimated from the macro dataset  $\chi^M$  (since the raw data used to form this factor  $q_{1t}$  are transformed to stationary), which we denote  $Q_{1t}$ . Since our emphasis is on  $h = 1$ , we write  $U_{Mt}$  instead of  $U_{Mt}(1)$ , and analogously for  $U_{Ft}$ , in order to simplify notation. Our *baseline* VAR is defined by the system that uses production as a measure of real activity, along with  $U_{Mt}$  and  $U_{Ft}$ , i.e.,  $\mathbf{X}_t = (U_{Mt}, ip_t, U_{Ft})'$ .

The bounds for the correlation and event constraints in each of these cases are set as follows. We start with the baseline system using  $Y_t = ip_t$  and set the  $\bar{c}$  and  $\bar{C}$  to be relatively unrestrictive, with  $\bar{c} = 0.03$  for the individual correlation, and  $\bar{C} = 0.24$  for the collective correlation. The latter value corresponds to an average value of approximately 0.14 for the root-mean-square-correlation  $\sqrt{\frac{1}{3}c(\boldsymbol{\beta})'c(\boldsymbol{\beta})}$ . This says that a lower bound of 3% absolute correlation between stock market returns and both types of uncertainty shocks is maintained, with an average absolute correlation of 14%. We set the parameters of the event constraints to  $\bar{k}_1 = 4.0$ ,  $\bar{k}_2 = 4.0$ , and  $\bar{k}_3 = 2$ . The  $\bar{k}_1$  and  $\bar{k}_2$  thresholds pertain to the financial uncertainty shocks in October of 1987 when Black Monday occurred, and during the months of the 2007-09 financial crisis. This imposes the constraint that these events were accompanied by large financial uncertainty shocks. The requirement that the shocks be at least four standard deviations larger than the mean is roughly guided by Bloom (2009). In his work, uncertainty shocks are calibrated from innovations to the VXO stock market volatility index. Bloom (2009) studies the dynamic effects of four standard deviation shocks to uncertainty. The  $\bar{k}_3$  threshold states that the identified real activity shock cannot be too positive in the Great Recession; specifically, we restrict the shock to be no larger than two standard deviations above its sample mean in the Great Recession months. In the additional cases section below, we assess the sensitivity of the findings to changing the values of the bounds. It is worth noting that, for the baseline case, the event constraints alone eliminate 99% of the solutions in  $\hat{\mathcal{B}}$ . When combined with the correlation constraints, we are left with a handful of solutions.

For the systems with other measures of  $Y_t = emp_t$  or  $Q_{1t}$ , the correlation bounds must be strengthened in order for the constraints to have a similar degree of restrictiveness. Specifically,

if for these cases we use the same  $\bar{c} = 0.03$  for the individual correlation and  $\bar{C} = 0.24$  for the collective correlation, more than three times as many solutions survive the correlation constraints (for the same event constraints) than for the baseline system that uses  $Y_t = ip_t$ . Thus it's clear that the correlation bounds parameters for the  $ip_t$  system have very little identifying power for the systems using  $emp_t$  and  $Q_{1t}$ . To make the degree of restrictiveness of these constraints comparable with the baseline case, we increase the individual correlation parameter to  $\bar{c} = 0.05$  for both the  $Y_t = emp_t$  or  $Q_{1t}$  systems, while setting collective correlation  $\bar{C} = 0.247$  for the  $emp_t$  system and  $\bar{C} = 0.244$  for the  $Q_{1t}$  system. The event constraint parameters  $\bar{k}$  are the same as for the baseline case.

The top panel of Figure 1 plots the estimated macro uncertainty  $U_{Mt}$  in standardized units along with the NBER recession dates. The horizontal bar corresponds to 1.65 standard deviation above unconditional mean of each series (which is standardized to zero). As is known from JLN, the macro uncertainty index is strongly countercyclical, and exhibits large spikes in the deepest recessions. The updated data  $U_{Mt}$  series shows much the same. Though  $U_{Mt}$  exceeds 1.65 standard deviations 48 times, they are clustered around the 1973-74 and 1981-82 recessions, as well as the Great Recession of 2007-09. Macroeconomic uncertainty is countercyclical and has a correlation of -0.65 with the 12-month moving-average of the growth in industrial production.

The middle panel of Figure 1 plots the financial uncertainty series  $U_{Ft}$  over time, which is new to this paper.  $U_{Ft}$  is a broad-based measure of time varying financial uncertainty using data from the bond market, stock market portfolio returns, and commodity markets. Hence, it is smoother than proxies such as VIX or any particular bond index. As seen from Figure 1,  $U_{Ft}$  is also countercyclical, though less so than  $U_{Mt}$ ; the correlation with industrial production is -0.39. The series often exhibits spikes around the times when  $U_{Mt}$  is high. However,  $U_{Ft}$  is more volatile and spikes more frequently outside of recessions, the most notable being the 1987 stock market crash. Though observations on  $U_{Ft}$  exceed the 1.65 standard deviation line 33 times, they are spread out in seven episodes, with the 2008 and 1987 episodes being the most pronounced. (The bottom panel of 1 plots the real activity uncertainty series  $U_{Rt}$  over time, discussed below.)

As is clear from Figure 1, both indicators of macro and financial uncertainty are serially correlated and hence predictable. They have comovements but also have independent variations as the correlation between them is only 0.58. However, this unconditional correlation cannot be given a structural interpretation. The heightened uncertainty measures can be endogenous responses to events that are expected to happen, but they can also be exogenous innovations. We use a VAR to capture the predictable variations, and then identify uncertainty shocks from the VAR residuals using the restrictions described in the previous section.

## 4.1 SVAR Estimates and Uncertainty Shocks

Several features of the VAR estimates are qualitatively similar for all measures of  $Y_t$ . Table 1 highlights some of these results. For the purposes of this table, we show estimates for the single max-C solution in (10). Panel A of this table shows that the sample correlation coefficients between  $Z_{1t}$  and  $\hat{e}_{Mt}$  and  $\hat{e}_{Ft}$ , and between  $Z_{2t}$  and  $\hat{e}_{Ft}$  are negative in each case, indicating that uncertainty shocks of both types tend to be high when stock market returns are low. Panel A also shows that the correlation between  $Z_{1t}$  and  $\hat{e}_{Yt}$ , and the correlation between  $Z_{2t}$  and  $\hat{e}_{Yt}$  and  $\hat{e}_{Mt}$  are all zero, which is true by construction of the algorithm and solution for **B**. Panel B shows that  $\sigma_{MM}$ ,  $\sigma_{YY}$ , and  $\sigma_{FF}$  are all non-zero for the max-C solution. In square brackets we report the range of values for these parameters across all solutions in the constrained set. The max-C solution is roughly in the middle of the range, which includes very small values close to zero all the way up to values several times higher. This in turn indicates the presence (for many solutions) of both macro and financial uncertainty shocks in the SVAR, as well as real activity shocks. Since both  $U_{Mt}$  and  $U_{Ft}$  are serially correlated, we should therefore find that  $Z_{1t}$  is correlated with lags of  $U_{Mt}$  and  $U_{Ft}$ , while  $Z_{2t}$  is correlated with lags of  $U_{Ft}$ . Results not reported confirm this is the case.

Figure 2 presents the time series of the standardized shocks ( $e_M, e_{ip}, e_F$ ) identified from the system with  $Y_t = ip_t$ , again for the max-C solution. All shocks display strong departures from normality with excess skewness and/or excess kurtosis. The largest of the  $e_{ip}$  shocks is recorded in 1980:04, followed by 1974:11, and 2005:09. There also appears to be a moderation in the volatility of the  $ip$  shocks in the post-1983 period. The largest macro uncertainty shock is in 1970:12, followed by the shock in 2008:10. The largest financial uncertainty shock is recorded in 1987:10 (Black Monday), followed by the shock in 2008:09 during the financial crisis. For  $e_F$ , the extreme but transitory nature of the 1987 stock market crash leads to a very large spike upward in  $e_F$  in the month of the crash, followed by a very large spike downward in the month following the crash as the market recovered strongly and quickly. While this episode magnifies the spike in  $e_F$  in 1987, it is largely orthogonal to real activity and macro uncertainty.

Observe that the large  $ip$  shock in 2005:09 is not associated with a contemporaneous spike in uncertainty, while there are several spikes in both types of uncertainty that do not coincide with spikes in  $e_{ip}$ . The next subsection uses impulse response functions to better understand the dynamic causal effects and propagating mechanisms of these shocks.

## 4.2 The Dynamic Effects of Uncertainty Shocks

Impulse response functions (IRFs) trace out the effects of counterfactual increases in the shocks. All plots show responses to one standard deviation changes in  $\epsilon_{jt}$  in the direction that leads to an *increase* in its own variable  $X_{jt}$ .

The left panel of Figure 3 shows in shaded areas the set of dynamic responses that satisfy the winnowing constraints for each variable in the SVAR to each structural shock for the baseline system with  $Y_t = ip_t$ . The dotted line shows the max-C solution. The right panel displays the analogous plots for systems that use  $emp_t$ , and the real activity index  $Q_{1t}$ . To avoid clutter, the max-C solutions are omitted from the right panel, but we comment on them below.

The figures show that positive shocks to *financial* uncertainty  $e_F$  (center plot, bottom row) lead to sharp declines in all three measures of real activity that persists for many months. All solutions that satisfy the identification restrictions have this pattern. These results lend support to the hypothesis that heightened financial uncertainty is an exogenous impulse that causes declines in real activity. However, there is little evidence that high financial uncertainty is a consequence of lower economic activity. Instead, exogenous (positive) shocks to real activity either increase financial uncertainty or have no clear affect on it.

Positive perturbations to  $e_{Ft}$  also cause  $U_{Mt}$  to increase sharply. However, there is less evidence that shocks to macro uncertainty have effects on financial uncertainty: the set of solutions show positive response of financial uncertainty for the baseline system with  $Y = ip$ , but the responses for the other two measures of real activity range from positive to zero to negative.

While we find no evidence that high financial uncertainty is a consequence of lower economic activity, the results for macro uncertainty are quite different. Both figures show that macro uncertainty falls sharply in response to positive real activity shocks when real activity is measured as  $ip_t$  or  $emp_t$ . Alternatively stated, negative real activity shocks increase macro uncertainty sharply. These endogenous movements in macro uncertainty persist for well over a year after the real activity shock. This result is strongly apparent in all the solutions of the identified sets for  $Y$  measured as production or employment, suggesting that higher macro uncertainty in recessions is a direct endogenous response to lower economic activity. The responses in the system using the real activity index  $Q_{1t}$  as a measure of  $Y$  are inconclusive, as the identified set in this case includes a wide range surrounding zero even though the max-C solution (not displayed) indicates that  $U_M$  falls sharply in response to a positive  $Q_1$  shock.

Finally, there is no evidence that the observed negative correlation between macro uncertainty and real activity is driven by causality running in the opposite direction. Indeed, the top middle panels of each figure show that positive macro uncertainty shocks often *increase* real activity in the short run, consistent with growth options theories discussed above. The exception again is the system with  $Y = Q_1$  where the identified set displays a wide range of responses. In all cases the max-C solution implies that real activity increases initially after a positive shock to macro uncertainty.

### 4.3 The Structural Shocks and Decomposition of Variance

In Figure 1 presented earlier, we find 1973-74, 1981-82, and 2007-2009 to be the three episodes of heightened macroeconomic uncertainty, defined as the periods when  $U_{Mt}$  is at least 1.65 standard deviations above its unconditional mean. We now look for the “large adverse” shocks in the systems  $(U_{Mt}, Y_t, U_{Ft})'$ , with  $Y_t = ip_t, emp_t, Q_{1t}$ . More precisely, we consider large positive uncertainty shocks and large negative real activity shocks.

For the max-C solution, the left panel of Figure 4 displays the date and size of shocks that are at least two standard deviations above the mean, estimated using the three different measures of  $Y_t$ . In view of the non-normality of the shocks, the figure also plots horizontal lines corresponding to three standard deviation of the unit shocks, which is used as the reference point for ‘large’. The lowest panel shows that, irrespective of the definition of  $Y_t$ , all SVARs identify big financial uncertainty shocks in October 1987 and in one or more months of 2008. Such solutions are selected as part of the identification scheme. The middle panel shows that large negative real activity shocks are in alignment with all post-war recessions with one exception: the negative real activity shock in 2005 is not immediately associated with a recession, but it could be the seed of the Great Recession that followed. It is known that the housing market led the 2007-2009 recession (e.g., see Favilukis, Ludvigson, and Van Nieuwerburgh (2015) for a discussion). Indeed, all 10 housing series in  $\mathcal{X}^M$  (most pertaining to housing starts and permits series) exhibit sharp declines starting in September 2005 and continuing through 2006, thereby leading the Great Recession. This suggests that the negative spike in real activity in 2005 was partly driven by the housing sector.

The left panel of Figure 4 shows that the dates of large increases in  $e_M$  are less clustered. They generally coincide with, or occur shortly after, the big real activity shocks and the financial uncertainty shocks. Observe that large macro uncertainty shocks occurred more frequently in the pre-1983 than the post 1983 sample, consistent with a Great Moderation occurring over the period ending in the Great Recession. However, increases of greater than three standard deviations for  $e_M$  appear only when real activity is measured by production in the SVAR, a point we return to below.

To give a sense of the historical importance of these shocks, we perform a decomposition of variance, given by the fraction of  $s$ -step-ahead forecast error variance attributable to each structural shock  $\epsilon_{Mt}$ ,  $\epsilon_{Yt}$ , and  $\epsilon_{Ft}$  for  $s = 1, s = 12, s = \infty$ . We also report the maximum fraction of forecast error variance over all VAR forecast horizons  $s$  that is attributable to each shock, denoted  $s = s_{max}$  in Table 2. The top panel of Table 2 reports these results for the max-C solutions in the systems with  $U_{Mt}$  and with  $Y_t = ip_t$  (left column),  $Y_t = emp_t$  (middle column), and  $Y_t = Q_{1t}$  (right column). In square brackets we report, for the  $s_{max}$  horizon, the range of values from lowest to highest across all solutions in the constrained set.

According to the top row, all three real activity shocks have sizable effects on macroeconomic uncertainty  $U_M$ , with shocks to production explaining up to 66% of the variation in  $U_M$ . But according to the bottom row, these same shocks have small effects on financial uncertainty  $U_F$ . At the same time, positive macro uncertainty shocks  $e_M$ , which increase rather than decrease real activity, explain a surprisingly large fraction of production (up to 69%), employment (up to 81%) and the real activity index (up to 85%). The ranges across all survived solutions tend to be wide, however. On the other hand, financial uncertainty shocks  $e_F$  have a small contribution to the one-step-ahead forecast error variance of all three measures of real activity, but their relative importance increases over time. Financial uncertainty shocks explain up to 40% of the forecast error variance in production, up to 44% of the forecast error variance in employment, and up to 40% of the forecast error variance in the real activity index. Financial uncertainty shocks  $e_F$  feed into  $U_M$ , and macroeconomic uncertainty shocks  $e_M$  also feed into  $U_F$ .

Regardless of which measure of real activity is used, we find that financial uncertainty is unlike macro uncertainty or real activity in that its variation is far more dominated by its own shocks. For example, in the system with  $ip$ ,  $e_F$  shocks explain 97% of the  $s = 1$  step-ahead forecast error variance in  $U_{Ft}$ , and 95% of the  $s = \infty$  step-ahead forecast error variance. In the systems with  $emp$  and  $Q_1$ ,  $e_F$  shocks explain 95% and 96%, respectively, of the  $s = 1$  step-ahead forecast error variance in  $U_{Ft}$ , and 90% and 78%, respectively, of the  $s = \infty$  step-ahead forecast error variance.

To summarize, in all three systems, real activity shocks  $e_Y$  have quantitatively large persistent negative effects on macro uncertainty  $U_M$ . In turn, macro uncertainty shocks  $e_M$  often have large positive impact effects on real activity measures  $Y$ . Financial uncertainty shocks  $e_F$  have smaller impact effects but larger long run effects that dampen real activity  $Y$ . Across all systems, the forecast error variance of financial uncertainty is the least affected by shocks other than its own, suggesting that  $U_F$  is quantitatively the most important exogenous impulse in the system.

## 5 Uncertainty in Real Activity $\mathbf{X}_t = (U_{Rt}, Y_t, U_{Ft})'$

The results discussed above suggest that the dynamic relationship between macro uncertainty and real activity can be quite different from the relation between financial uncertainty and real activity. However, given the composition of our data  $\chi^M$ , macroeconomic uncertainty itself can be due to uncertainty in real activity variables such as output and unemployment, to price variables, and to financial market variables. The theoretical uncertainty literature has focused on modeling exogenous uncertainty shocks that arise specifically in measures of real economic fundamentals, rather than in prices or financial markets. To better evaluate the implications of



these theoretical models, we consider systems that isolate uncertainty about real activity using the  $U_{Rt}$  sub-index that more closely corresponds to the theoretical literature.

We isolate the real activity components of macro uncertainty by aggregating the individual uncertainty estimates over the 73 real activity variables in the macro dataset  $\mathcal{X}^M$ . The one-period ahead uncertainty in real activity, denoted  $U_{Rt}$ , is shown in the bottom panel of Figure 1. This series, like  $U_{Mt}$ , is countercyclical though somewhat less so, having a correlation of -0.50 with industrial production (as compared to -0.66 for  $U_{Mt}$ ). At first glance,  $U_{Rt}$  appears to fluctuate in a manner similar to macroeconomic uncertainty  $U_{Mt}$ . The two series have a correlation of 0.71 and exhibit some overlapping spikes. But  $U_{Rt}$  and  $U_{Mt}$  also display notable independent variation. The bottom panel of Figure 1 shows that there are 43 observations of  $U_{Rt}$  that are at least 1.65 standard deviations above its mean. These can be organized into five episodes: 1965, 1970, 1975, 1982-83, and 2007. By contrast,  $U_{Mt}$  in the top panel of Figure 1 only exhibits three such episodes. Observe that the  $U_{Rt}$  series exhibits several spikes before 1970 that are not accompanied by spikes in  $U_{Mt}$ .

The bounds for the correlation and event constraints for these systems are set as follows. An inspection of the solution sets for systems that use  $U_{Rt}$  in place of  $U_{Mt}$  reveals that positive financial uncertainty  $e_{Ft}$  shocks are larger in the financial crisis/Great Recession for systems that use  $U_{Rt}$ , both in terms of the average increase in the periods of this episode as well as the magnitudes of the shocks in the upper quantiles. This implies that the bounds on the behavior of financial uncertainty shocks in the financial crisis need to be strengthened to be made comparable to the case that uses  $U_{Mt}$ . Moreover, in the system with  $U_{Rt}$  an individual correlation constraint of  $\bar{c} = 3\%$  or even  $5\%$  is almost entirely unrestrictive. To ensure that the restrictiveness of the bounds in this case is comparable to that of the systems with  $U_{Mt}$  we therefore set the bounds  $k_1$  and  $k_2$  on the  $e_{Ft}$  shock behavior to be 4.3 (rather than 4) standard deviations for systems that use  $emp_t$  and  $Q_{1t}$ , and 4.1 for the system that uses  $ip_t$ . We set the individual correlation parameter to  $\bar{c} = 0.07$  and the collective correlation parameter to  $\bar{C} = 0.24$  in all cases.

Given the distinctive patterns in the time series behavior of  $U_{Rt}$  and  $U_{Mt}$ , one might expect to find different dynamic relationships with the other variables in our systems when  $U_{Mt}$  is replaced by  $U_{Rt}$ . However, the impulse responses functions are qualitatively similar to systems studied above that use broad-based macro uncertainty. The sets of responses that satisfy our winnowing constraints are displayed in Figure 5. The solid line for the  $Y_t = ip_t$  system indicates that the set in that case is a singleton. The figure shows that, for all solutions that satisfy the constraints and no matter which measure of real activity is used, (i) positive shocks to real activity measures unambiguously cause sharp declines in real economic uncertainty  $U_{Rt}$  so that negative shocks cause sharp increases in real economic uncertainty; (ii) positive real activity uncertainty shocks  $e_{Rt}$  do not cause declines in real activity measures; indeed the opposite

is unambiguously true; (iii) positive financial uncertainty shocks  $e_{Ft}$  lead to declines in real activity measures that are steep and persistent, and (iv) there is little evidence that high (low) financial uncertainty is caused by negative (positive) real activity shocks; the sets of IRFs are either positive or surround zero for all systems. Thus the identified sets present an even clearer picture of the dynamic causal relationships in the systems with  $\mathbf{X}_t = (U_{Rt}, Y_t, U_{Ft})'$  than they do in the systems with  $\mathbf{X}_t = (U_{Mt}, Y_t, U_{Ft})'$ .

The right panel of Figure 4 plots the large adverse structural shocks for the max-C solutions identified from the systems  $(U_{Rt}, Y_t, U_{Ft})'$  for  $Y_t = ip_t, emp_t, Q_{1t}$ . The topmost right panel shows that the real uncertainty shock  $e_{Rt}$  exhibits spikes in excess of three standard deviations during the Great Recession only for the system in which  $Y_t = ip_t$ . Moreover, for the other two systems in which  $Y_t = emp_t$  or  $Q_{1t}$ , there is not a spike that exceeds even two standard deviations above its mean, despite the fact that  $U_{Rt}$  itself exhibits a large spike in the Great Recession. These episodes serve to reinforce the conclusion from the IRFs that the heightened real economic uncertainty in recessions is often endogenous response to other shocks, rather than an exogenous impulse. Even though there were many large spikes in real uncertainty shocks  $e_{Rt}$  pre-1983, there have been fewer large adverse shocks to real economic uncertainty since 1983, a period that coincides with the so-called Great Moderation.

To complete the analysis, we present variance decompositions for the system  $(U_{Rt}, Y_t, U_{Ft})'$ , with three measures of real activity  $Y_t = ip_t, emp_t, Q_{1t}$ . These results, presented in the bottom panel of Table 2, share some similarities with the systems that use macro uncertainty  $U_{Mt}$  shown in the top panel, but there are at least two distinctions. First, financial uncertainty shocks decrease real activity and explain larger fractions of the forecast error variance in all three measures of real activity at long horizons. The ranges for these numbers at the  $s = s_{max}$  horizon across all solutions in the winnowed set are also relatively narrow. Second, compared to systems that use  $U_{Mt}$ , smaller fractions of the forecast error variance in  $U_{Rt}$  are explained by its own shocks. This is because, in these systems, shocks to all three measures of real activity have larger effects on real activity uncertainty than they do on macro uncertainty. This reinforces the point that countercyclical increases in real economic uncertainty are often well characterized as endogenous responses to declines in real activity, rather than exogenous impulses driving real activity downward.

## 6 Additional Cases

This section presents results for a number of additional cases.

## 6.1 Alternative Bounds

Our base cases imposes two types of winnowing constraints, the event constraints and the correlation constraints. If either of these were not helpful in narrowing the solution sets, then failure to impose one or the other would not have a significant affect on the results. Figure 6 shows the baseline IRF when  $\bar{k}$  is changed from (4,4,2) to (2,2,3) with  $(\bar{c}, \bar{C})$  held fixed at (0.03, 0.24). This means that shocks of less extreme magnitudes are admitted into the solution set. This leaves about seven times more solutions (879 compared to 108). The right panel shows the sets of IRFs when the correlation constraint bounds  $(\bar{c}, \bar{C})$  are changed from (0.03, 0.24) to (0.015, 0.12) with  $\bar{k}$  held fixed at (4,4,2). This means the synthetic variables are allowed to be less correlated with the shocks. This leaves about three times more solutions (407 compared to 108 under the baseline parameterization). In both cases, many of the additional solutions fail to show any large spike in financial uncertainty in the financial crisis or a non-negligible correlation between either type of uncertainty shock and stock market returns.

Figure 6 shows that the identified sets of IRFs in both panels are noticeably wider indicating that each type of constraint contributes to identification. The left panel shows that all solutions imply that financial uncertainty shocks drive  $ip_t$  down eventually, though the effect is smaller than the base case. Positive real activity shocks either drive down or have no effect on macro uncertainty. While the responses of real activity to macro uncertainty shocks are wide, a macro uncertainty shock has an inconclusive effect on macro uncertainty itself. The right panel shows less restrictive correlation bounds also yield a wider range of impulse responses. The noticeable difference compared to the baseline case concerns the effects financial uncertainty shocks on real activity, which are not well determined. This suggests that the correlation constraints matter most for the effect of financial uncertainty shocks. Solutions for which financial uncertainty shocks have a very low correlation with the stock market are unlikely to signify an important role for financial uncertainty in business cycle downturns. The effects of  $e_{Ft}$  shocks on the other variables in the system are insensitive to changes in either type of bound. Taken together, the results in two panels demonstrate the importance and identifying power of both types of constraints for drawing clear conclusions about the dynamic causal effects in the system.

## 6.2 Validity of Recursive Identification Restrictions

The econometric model permits us to test whether a recursive structure is supported by the data. Specifically, the assumptions in our event and correlation constraints do not rule out the possibility of a recursive structure, so that if such a structure is consistent with the data, our identifying restrictions are free to recover it. With three variables in the SVAR, there are six possible recursive orderings corresponding to six different  $3 \times 1$  vectors of elements of  $\mathbf{B}$  that must be jointly zero. It is straightforward to assess whether our identified solutions are consistent

with a recursive structure by examining the distribution of solutions in the constrained set for four elements of the  $\mathbf{B}$  matrix:  $\hat{B}_{YF}$ ,  $\hat{B}_{YM}$ ,  $\hat{B}_{MY}$ , and  $\hat{B}_{MF}$ . None of the distributions contain any values near zero. The minimum absolute values in each case are .0005, .0037, .0066, and .0016, respectively, which are all bounded away from zero. The implication is that the recursive structure is inconsistent with any recursive ordering across all solutions in the identified set.

What happens to the dynamic responses when we nevertheless impose restrictions based on recursive identification (and freely estimate the rest of the parameters)? With these recursive restrictions the SVAR is point-identified so no winnowing constraints are needed. Of course, there are many possible recursive orderings, and inevitably, the estimated IRFs differ in some ways across these cases. However, the dynamic responses under recursive identification have one common feature that is invariant to the ordering. Results available on request show that, no matter which ordering is assumed in the recursive structure, macro uncertainty shocks appear to cause a sharp decline in real activity, much like financial uncertainty shocks, while positive real activity shocks have little effect on macro uncertainty in the short run and if anything increase it in the long run, as shown in the figure. This is in stark contrast to the results from our identification scheme, which is capable of recovering a recursive structure if it were true. But we fail to find such a structure. These results show that imposing a structure that prohibits contemporaneous feedback may spuriously suggest that macro uncertainty shocks are a cause of declines in real activity, rather than an endogenous response. The finding underscores the challenges of relying on convenient timing assumptions to sort out cause and effect in the relationship between uncertainty and real activity.

### 6.3 System Estimation Results

This section reports the results of estimating the model using the full system approach described above. We estimate a four variable system in  $(\mathbf{X}_t, S_t)'$  where  $S_t$  is measured as the return on the CRSP value-weighted stock market index excess return. The bounds for the winnowing constraints are the same as for the subsystem analysis for the same variables. The left panel of Figure 7 presents the set identified IRFs for the full system estimation. The figure shows that the results are qualitatively very similar to the subsystem case. As for that case, positive shocks to financial uncertainty drive down all measures of real activity sharply and persistently, but there is no evidence that positive shocks to macro uncertainty decrease real activity. Positive shocks to production and employment clearly drive down macro uncertainty, though like the subsystem analysis the results are inconclusive in this system when real activity is measured by the index  $Q_{1t}$ .

As discussed above, the subsystem exclusion restriction for  $S_t$  places overidentifying restrictions on the full system estimation. A simple way to evaluate this restriction is to compare the

impulse response functions estimated for the three variable subsystem for  $\mathbf{X}_t = (U_{Mt}, Y_t, U_{Ft})'$ , with those from the larger system that includes  $S_t$  but does not restrict the coefficients of  $S_{t-j}$  in the equations for  $\mathbf{X}_t$  to zero, for  $j \geq 1$ . The right panel of Figure 7 presents the sets of identified impulse responses that satisfy the constraints in each case, overlaid on one another. The identified sets lie almost on top of each other, indicating that the responses are little different. Indeed, the coefficients on lags of  $S_t$  appear to be close to zero in all three  $\mathbf{X}_t$  equations. The data thus appear qualitatively consistent with the assumption that stock returns can be excluded from the VAR for  $\mathbf{X}_t$ .

## 7 Conclusion

A growing body of research establishes uncertainty as a feature of deep recessions but leaves open two key questions: is uncertainty primarily a source of business cycle fluctuations or an endogenous response to them? And where does uncertainty originate? There is no theoretical consensus on the question of whether uncertainty is a cause or a consequence of declines in economic activity. In most theories, it is modeled either as a cause or an effect but not both, underscoring the extent to which the question is fundamentally an empirical matter.

The objective of this paper is to address both questions econometrically using small-scale structural VARs that are general enough to nest a range of theoretical possibilities in empirical tests. Commonly used recursive identification schemes cannot achieve this objective, since by construction they rule out the possibility that uncertainty and real activity could influence one another contemporaneously. The econometric model employed in this paper nests the recursive identification scheme, and we find that it is strongly rejected by the data.

To identify dynamic causal effects, this paper takes an alternative identification approach that imposes economic assumptions about the behavior of the structural shocks to allow sets of solutions to be identified. In addition, our empirical analysis explicitly distinguishes macro uncertainty and uncertainty about real activity from financial uncertainty, thereby allowing us to shed light on the origins of uncertainty shocks that drive real activity lower, to the extent that any of them do. The econometric framework permits uncertainty to be an exogenous source of business cycle fluctuations, or an endogenous response to them, or any combination of the two, without restricting the timing of these relationships. The results from these estimations show that sharply higher uncertainty about real economic activity in recessions is more likely to be an endogenous response to business cycle fluctuations, while uncertainty about financial markets is a likely source of them. Exogenous declines in economic activity have quantitatively large effects that drive real economic uncertainty endogenously higher. Financial uncertainty, by contrast, is dominated by its own shocks, implying that it is primarily an exogenous impulse vis-a-vis real activity and macro uncertainty.

## References

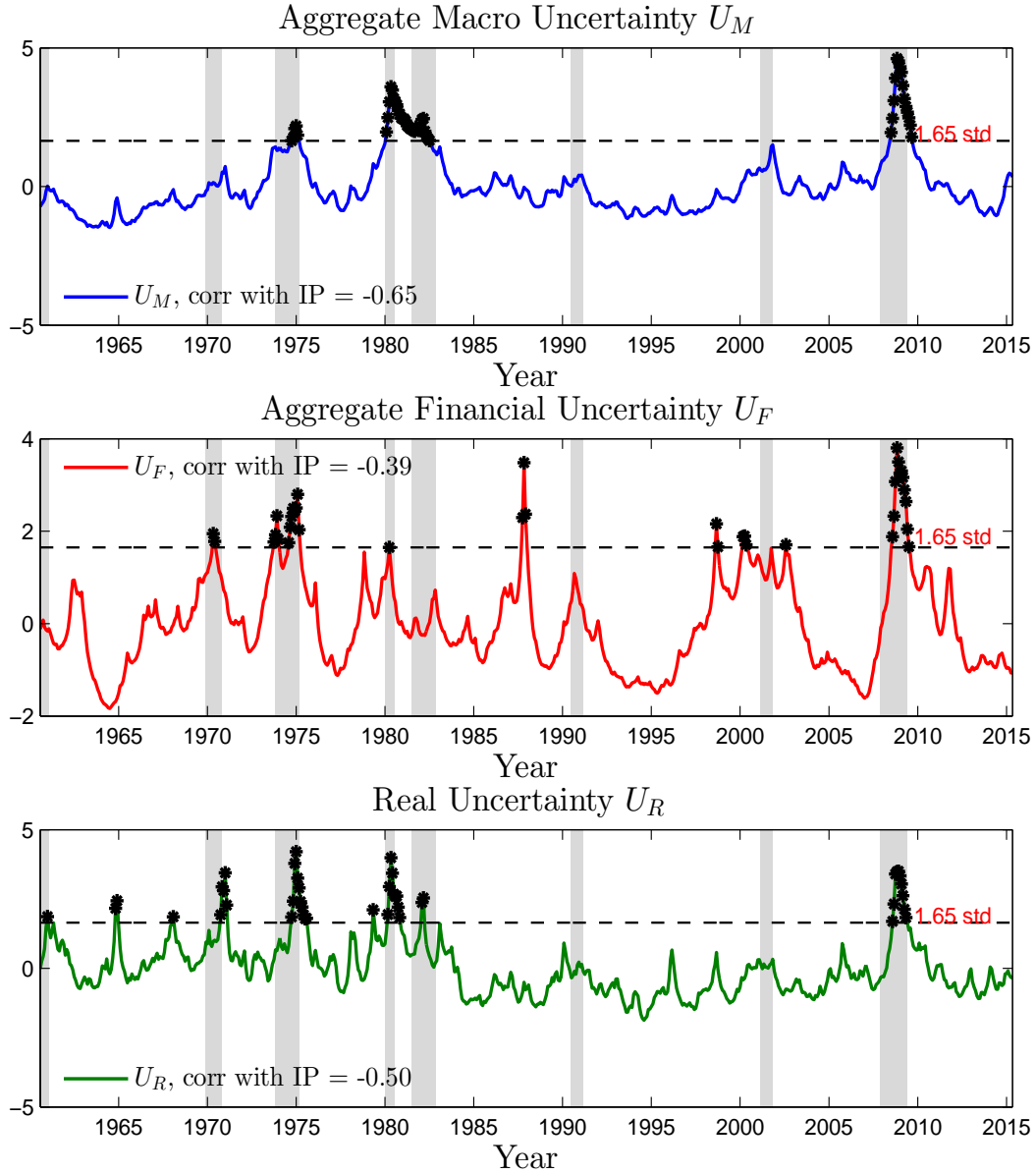
- AI, H., K. LI, AND F. YANG (2015): “Financial Intermediation and Capital Misallocation,” Unpublished manuscript, University of Minnesota.
- ARELLANO, C., Y. BAI, AND P. KEHOE (2011): “Financial Markets and Fluctuations in Uncertainty,” Federal Reserve Bank of Minneapolis Research Department Staff Report 466.
- BACHMANN, R., S. ELSTNER, AND E. R. SIMS (2013): “Uncertainty and Economic Activity: Evidence from Business Survey Data,” *American Economic Journal: Macroeconomics*, 5(2), 217–49.
- BACHMANN, R., AND G. MOSCARINI (2011): “Business Cycles and Endogenous Uncertainty,” Discussion paper, Society for Economic Dynamics.
- BAKER, S. R., AND N. BLOOM (2013): “Does Uncertainty Reduce Growth? Using Disasters as Natural Experiments,” National Bureau of Economic Research Working Paper 19475.
- BAKER, S. R., N. BLOOM, AND S. J. DAVIS (2013): “Measuring Economic Policy Uncertainty,” <http://www.policyuncertainty.com/media/BakerBloomDavis.pdf>.
- BANSAL, R., AND A. YARON (2004): “Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles,” *Journal of Finance*, 59(4), 1481–1509.
- BAR-ILAN, A., AND W. C. STRANGE (1996): “Investment lags,” *The American Economic Review*, 86(3), 610–622.
- BASU, S., AND B. BUNDICK (2012): “Uncertainty Shocks in a Model of Effective Demand,” <http://fmwww.bc.edu/ec-p/wp774.pdf>.
- BEKAERT, G., M. HOEROVA, AND M. L. DUCA (2013): “Risk, Uncertainty and Monetary Policy,” *Journal of Monetary Economics*, 60(7), 771–88.
- BERGER, D., I. DEW-BECKER, AND S. GIGLIO (2016): “Contractionary volatility or volatile contractions?,” Unpublished paper, Northwestern University.
- BERNANKE, B. S. (1983): “Irreversibility, Uncertainty, and Cyclical Investment,” *The Quarterly Journal of Economics*, 98(1), 85–106.
- BIANCHI, F., C. ILUT, AND M. SCHNEIDER (2014): “Uncertainty Shocks, Asset Supply and Pricing Over the Business Cycle,” Unpublished paper, Duke University.
- BLOOM, N. (2009): “The Impact of Uncertainty Shocks,” *Econometrica*, 77(3), 623–85.
- (2014): “Fluctuations in Uncertainty,” *The Journal of Economic Perspectives*, pp. 153–175.
- BLOOM, N., M. FLOETOTTO, N. JAIMOVICH, I. SAPORTA-EKSTEN, AND S. J. TERRY (2012): “Really Uncertain Business Cycles,” National Bureau of Economic Research Working Paper 18245.
- FAJGELBAUM, P., E. SCHAAL, AND M. TASCHEREAU-DUMOUCHEL (2014): “Uncertainty traps,” Discussion paper, National Bureau of Economic Research Working Paper No. W19973.
- FAVILUKIS, J., S. C. LUDVIGSON, AND S. VAN NIEUWERBURGH (2015): “The Macroeconomic Effects of Housing Wealth, Housing Finance and Limited Risk Sharing in General Equilibrium,” *Journal of Political Economy*, forthcoming.

- FERNÁNDEZ-VILLAVERDE, J., J. F. R.-R. PABLO GUERRÓN-QUINTANA, AND M. URIBE (2011): “Risk Matters: The Real Effects of Volatility Shocks,” *American Economic Review*, 6(101), 2530–61.
- FOSTEL, A., AND J. GEANAKOPOLOS (2012): “Why does bad news increase volatility and decrease leverage?,” *Journal of Economic Theory*, 147(2), 501–525.
- GABAIX, X., AND M. MAGGIORI (2013): “International Liquidity and Exchange Rate Dynamics,” Unpublished paper, New York University.
- GILCHRIST, S., J. W. SIM, AND E. ZAKRAJSEK (2010): “Uncertainty, Financial Frictions, and Investment Dynamics,” Society for Economic Dynamics 2010 Meeting Papers No.1285.
- GOMES, J. F., AND L. SCHMID (2016): “Equilibrium Asset Pricing with Leverage and Default,” <http://finance.wharton.upenn.edu/~gomesj/Research/CreditMarkets.pdf>.
- GOURIO, F. (2012): “Disaster Risk and Business Cycles,” *American Economic Review*, 102(6).
- GREENWALD, D., M. LETTAU, AND S. C. LUDVIGSON (2014): “Origins of Stock Market Fluctuations,” National Bureau of Economic Research Working Paper No. 19818.
- HAMILTON, J. D. (2003): “What is an oil shock?,” *Journal of econometrics*, 113(2), 363–398.
- HE, Z., AND A. KRISHNAMURTHY (2013): “Intermediary Asset Pricing,” *American Economic Review*, 103(2), 732–770.
- JURADO, K., S. C. LUDVIGSON, AND S. NG (2015): “Measuring Uncertainty,” *The American Economic Review*, 105(3), 117–1216.
- KILIAN, L. (2008): “The Economic Effects of Energy Price Shocks,” *Journal of Economic Literature*, 46(4), 871–909.
- KOZAK, S., AND S. SANTOSH (2014): “Linking Cross-Sectional and Aggregate Expected Returns,” Available at SSRN 2518256.
- KRAFT, H., E. S. SCHWARTZ, AND F. WEISS (2013): “Growth options and firm valuation,” Discussion paper, NBER Working Paper w18836.
- LEDUC, S., AND Z. LIU (2012): “Uncertainty Shocks are Aggregate Demand Shocks,” Federal Reserve Bank of San Francisco, Working Paper 2012-10.
- LETTAU, M., AND S. C. LUDVIGSON (2013): “Shocks and Crashes,” in *National Bureau of Economics Research Macroeconomics Annual: 2013*, ed. by J. Parker, and M. Woodford, vol. 28, pp. 293–354. MIT Press, Cambridge and London.
- LUDVIGSON, S. C., S. MA, AND S. NG (2016): “Shock Restricted Structural Vector-Autoregressions,” Manuscript, New York University.
- LUDVIGSON, S. C., AND S. NG (2007): “The Empirical Risk-Return Relation: A Factor Analysis Approach,” *Journal of Financial Economics*, 83(1), 171–222.
- MCCRACKEN, M. W., AND S. NG (2016): “FRED-MD: A Monthly Database for Macroeconomic Research,” *Journal of Business and Economic Statistics*, forthcoming, <http://www.columbia.edu/~sn2294/papers/freddata.pdf>.
- MCDONALD, R., AND D. SIEGEL (1986): “The Value of Waiting to Invest,” *The Quarterly Journal of Economics*, 101(4), 707–728.

- MERTENS, K., AND M. O. RAVN (2013): “The dynamic effects of personal and corporate income tax changes in the United States,” *The American Economic Review*, 103(4), 1212–1247.
- MUIR, T. (2014): “Financial Crises and Risk Premia,” Unpublished paper, Yale University.
- NAKAMURA, E., D. SERGEYEV, AND J. STEINSSON (2012): “Growth-Rate and Uncertainty Shocks in Consumption: Cross-Country Evidence,” National Bureau of Economic Research Working Paper 18128.
- NG, S., AND J. H. WRIGHT (2013): “Facts and Challenges from the Great Recession for Forecasting and Macroeconomic Modeling,” *Journal of Economic Literature*, 51(4), 1120–1154.
- OLEA, J. L. M., J. H. STOCK, AND M. D. WATSON (2015): “Inference in Structural VARs with External Instruments,” Unpublished manuscript, Harvard University.
- PASTOR, L., AND P. VERONESI (2006): “Was There A Nasdaq Bubble in the Late 1990s?,” *Journal of Financial Economics*, 81, 61–100.
- PÁSTOR, L., AND P. VERONESI (2013): “Political uncertainty and risk premia,” *Journal of Financial Economics*, 110(3), 520–545.
- SEGAL, G., I. SHALIASTOVICH, AND A. YARON (2015): “Good and bad uncertainty: Macroeconomic and financial market implications,” *Journal of Financial Economics*, 117(2), 369–397.
- STOCK, J. H., AND M. W. WATSON (2008): “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](http://www.nber.org/minicourse_2008.html)*.
- (2012): “Disentangling the Channels of the 2007-09 Recession,” *Brookings Papers on Economic Activity*, p. 81.
- VAN NIEUWERBURGH, S., AND L. VELDKAMP (2006): “Learning asymmetries in real business cycles,” *Journal of monetary Economics*, 53(4), 753–772.
- WACHTER, J. (2013): “Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Activity?,” *Journal of Finance*, 68, 987–1035.

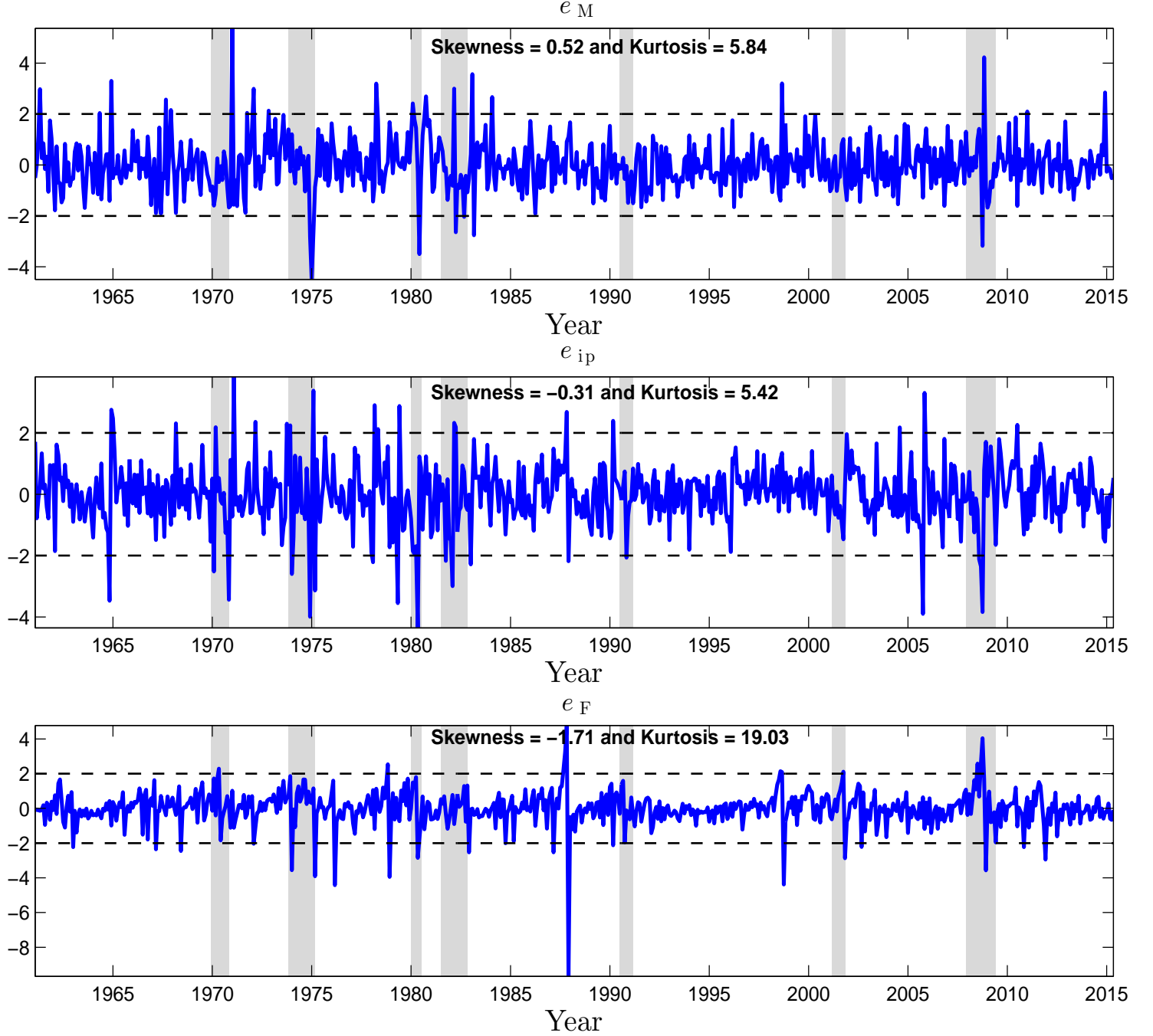


**Figure 1: Macro, Financial and Real Uncertainty Over Time**



The panels plot the time series of macro uncertainty  $U_M$ , financial uncertainty  $U_F$ , and real activity uncertainty  $U_R$ , expressed in standardized units. Shaded areas correspond to NBER recession dates. The horizontal line corresponds to 1.65 standard deviations above the unconditional mean of each series (which has been normalized to zero); the black dots are months when uncertainty is at least 1.65 standard deviations above the mean. Correlations with the 12-month moving average of IP growth are reported. The data span the period 1960:07 to 2015:04.

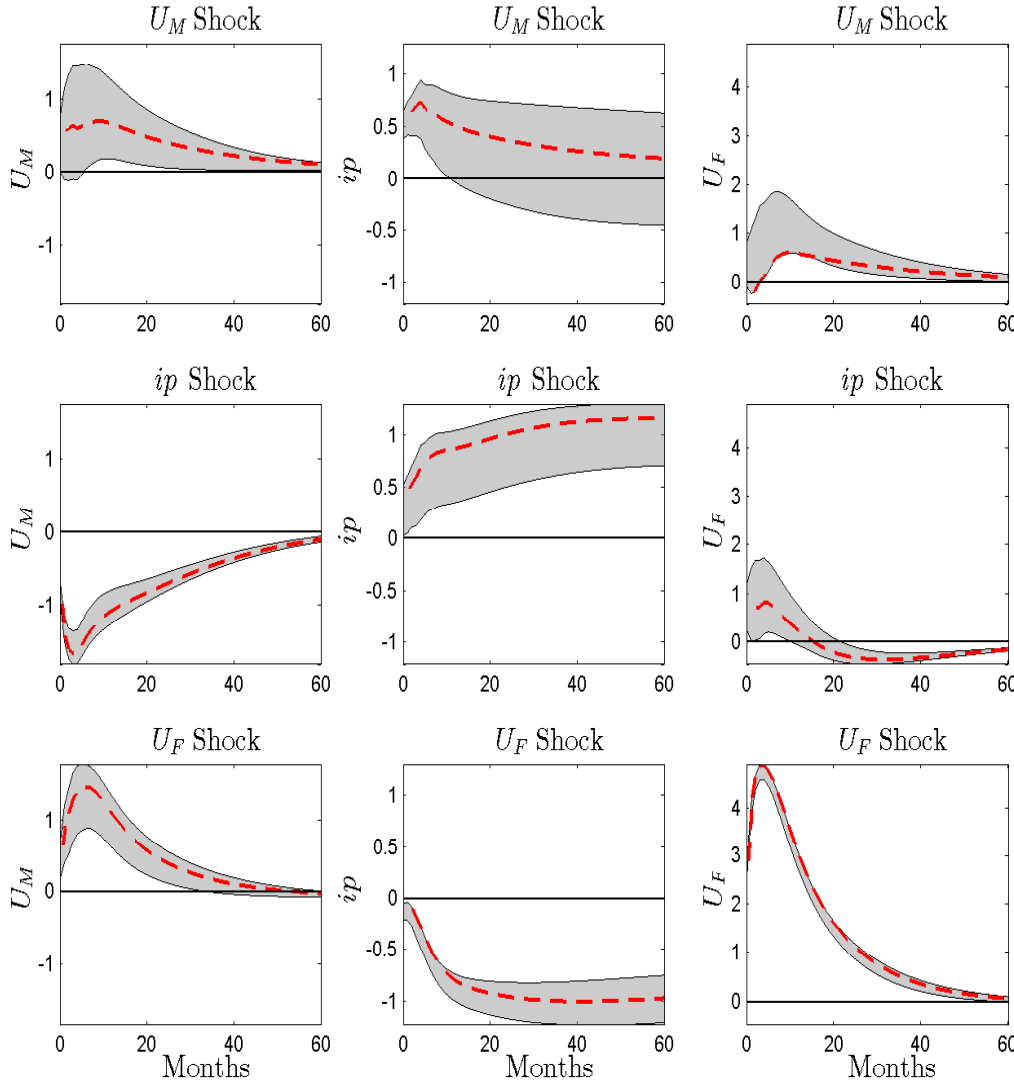
**Figure 2: Time Series of  $e$  Shock from SVAR  $(U_M, ip, U_F)'$**



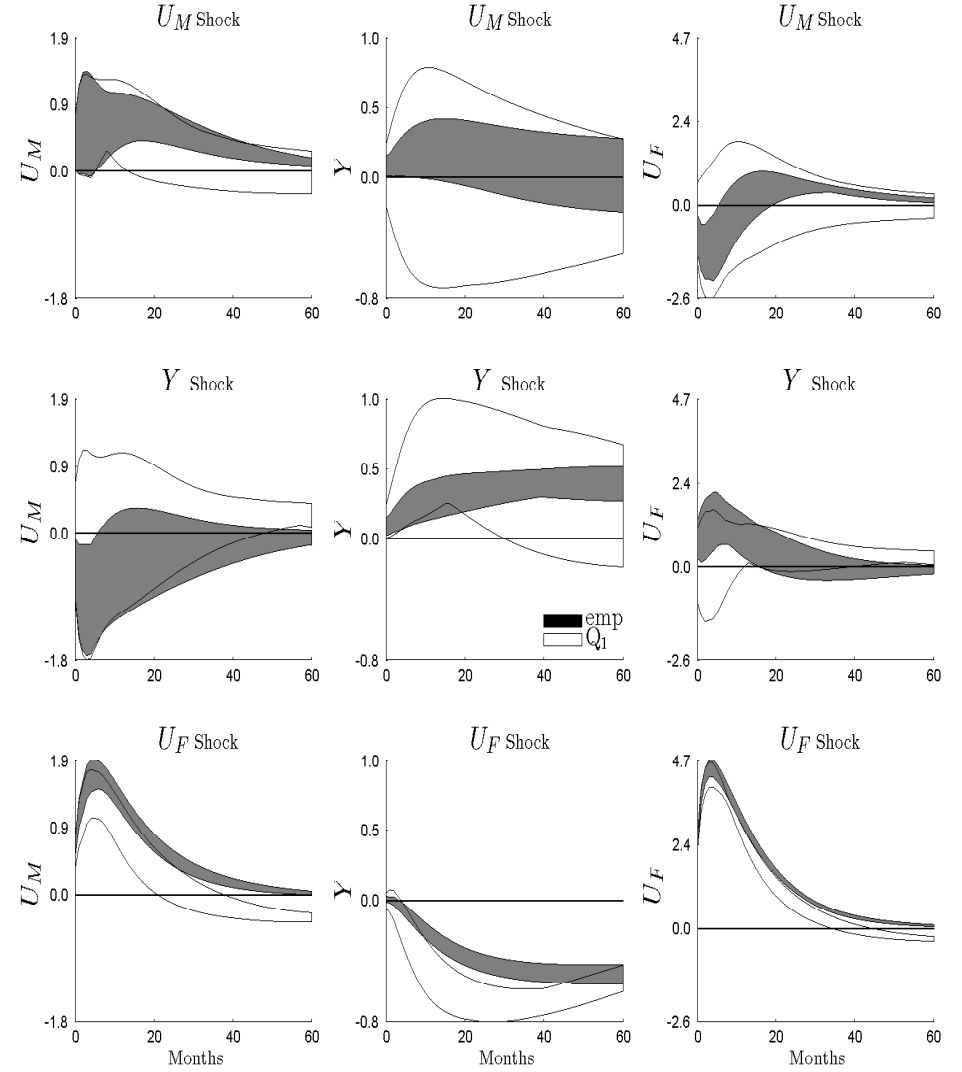
The horizontal line corresponds to 3 standard deviations above/below the unconditional mean of each series. The shocks  $e = B^{-1}\eta_t$  for max-C solution are reported, where  $\eta_t$  is the residual from VAR(6) of  $(U_M, ip, U_F)'$  and  $B = A^{-1}\Sigma^{\frac{1}{2}}$ . The sample spans the period 1960:07 to 2015:04.

**Figure 3: SVAR  $(U_M, Y, U_F)'$**

**IRFs of SVAR  $(U_M, ip, U_F)'$**

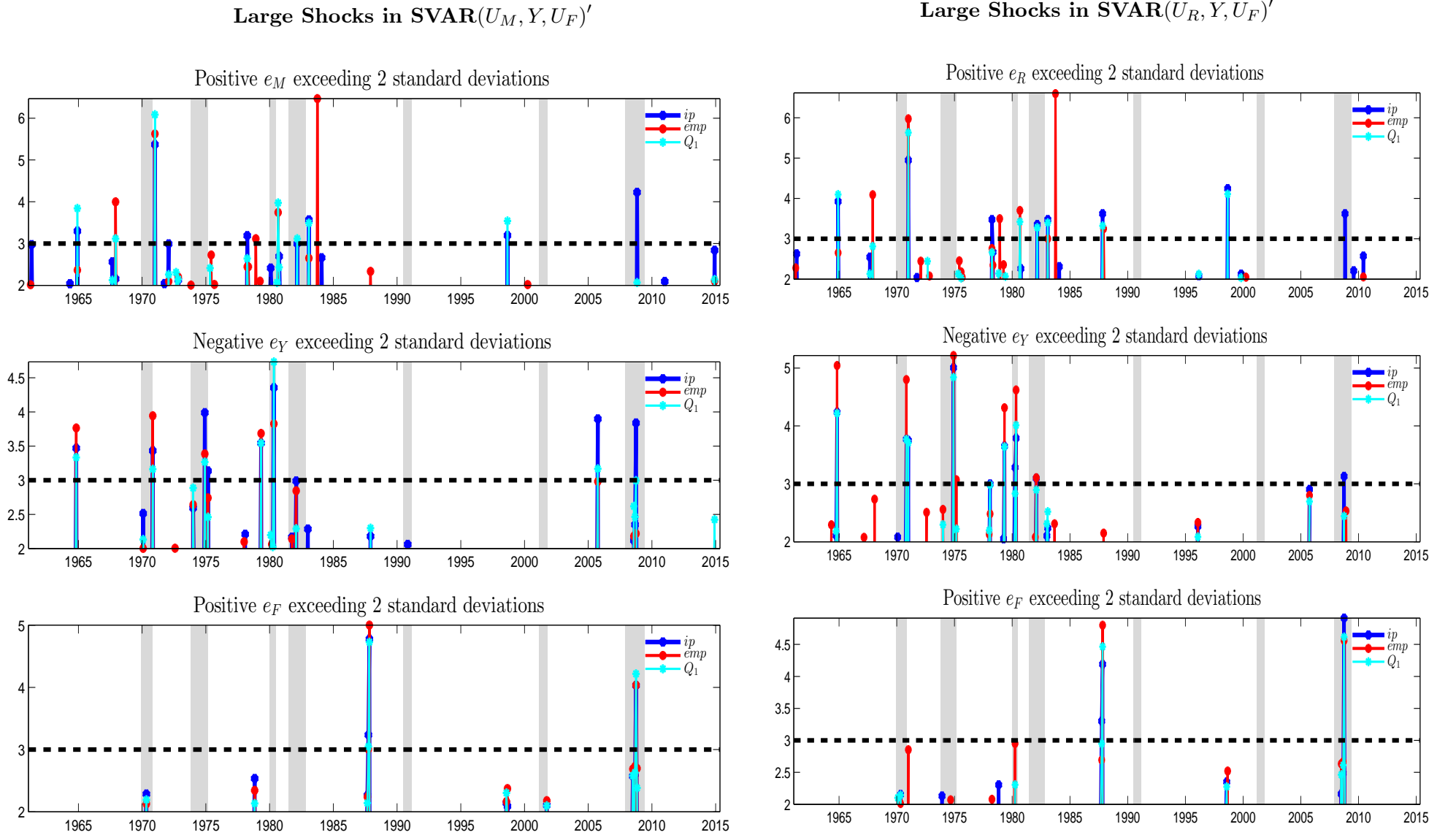


**IRFs of SVAR  $(U_M, Y, U_F)'$**



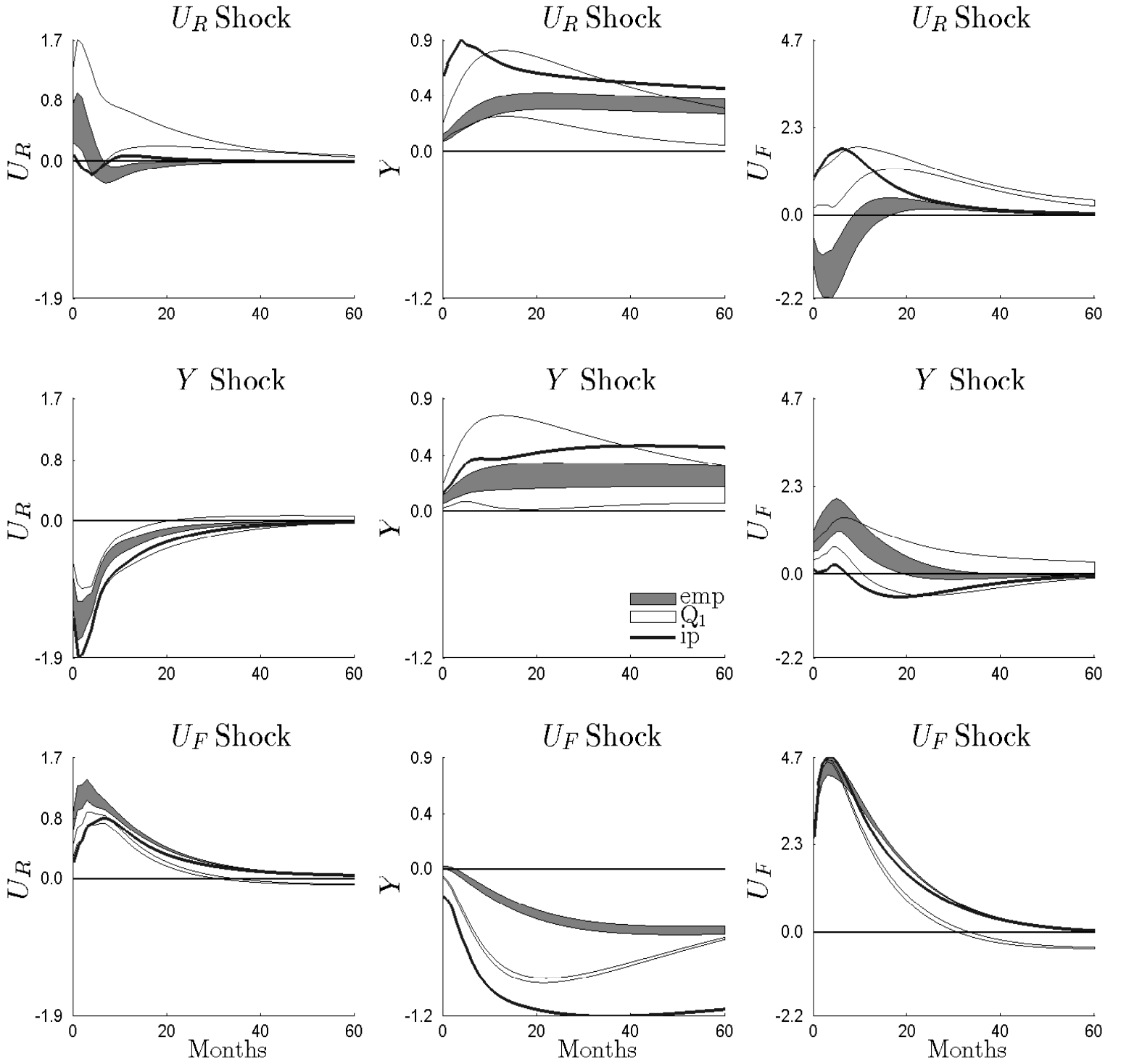
The left panel reports the IRFs of SVAR  $(U_M, ip, U_F)'$ . The dashed line is the max-C solution. The right panel reports the IRFs of SVAR  $(U_M, Y, U_F)'$  where  $Y = Q_1, emp$ . The shaded areas represent sets of solutions that satisfy the correlation and event constraints. Responses to positive one standard deviation shocks are reported in percentage points. The sample spans the period 1960:07 to 2015:04.

Figure 4: Large Shocks



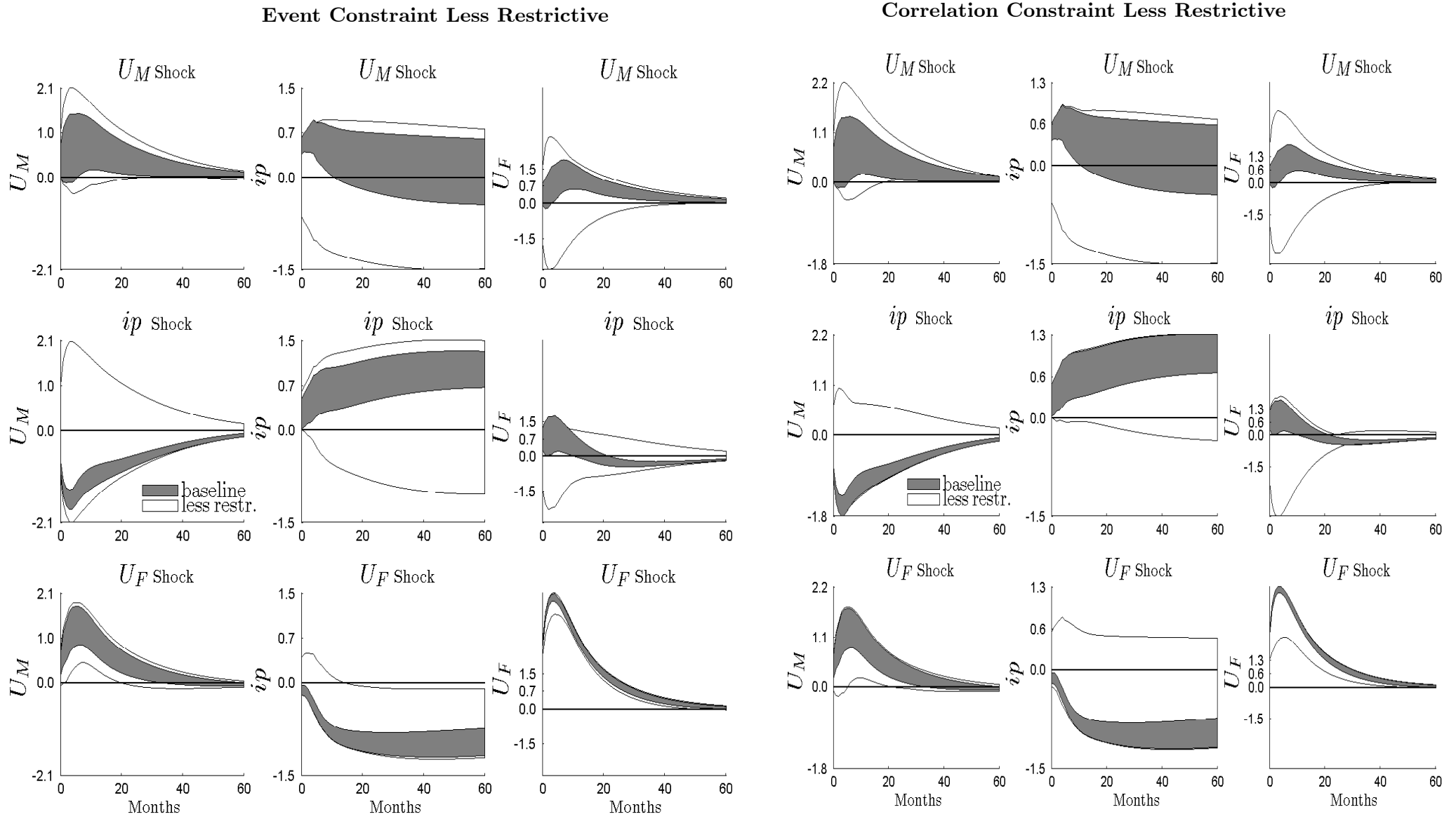
For the max-C solution, the figure exhibits shocks that are at least 2 standard deviations above the unconditional mean for  $e_M$  or  $e_R$  and  $e_F$  and below for  $e_Y$  for three cases where  $Y = ip, emp, Q_1$ . The shocks  $e_t = B^{-1}\eta_t$  are reported, where  $\eta_t$  is the residual from VAR(6) and  $B = A^{-1}\Sigma^{\frac{1}{2}}$ . The horizontal line corresponds to 3 standard deviations shocks. The sample spans the period 1960:07 to 2015:04.

**Figure 5: IRFs of SVAR  $(U_R, Y, U_F)'$**



The shaded areas represent sets of solutions that satisfy the correlation and event constraints. Responses to positive one standard deviation shocks are reported in percentage points. The sample spans the period 1960:07 to 2015:04.

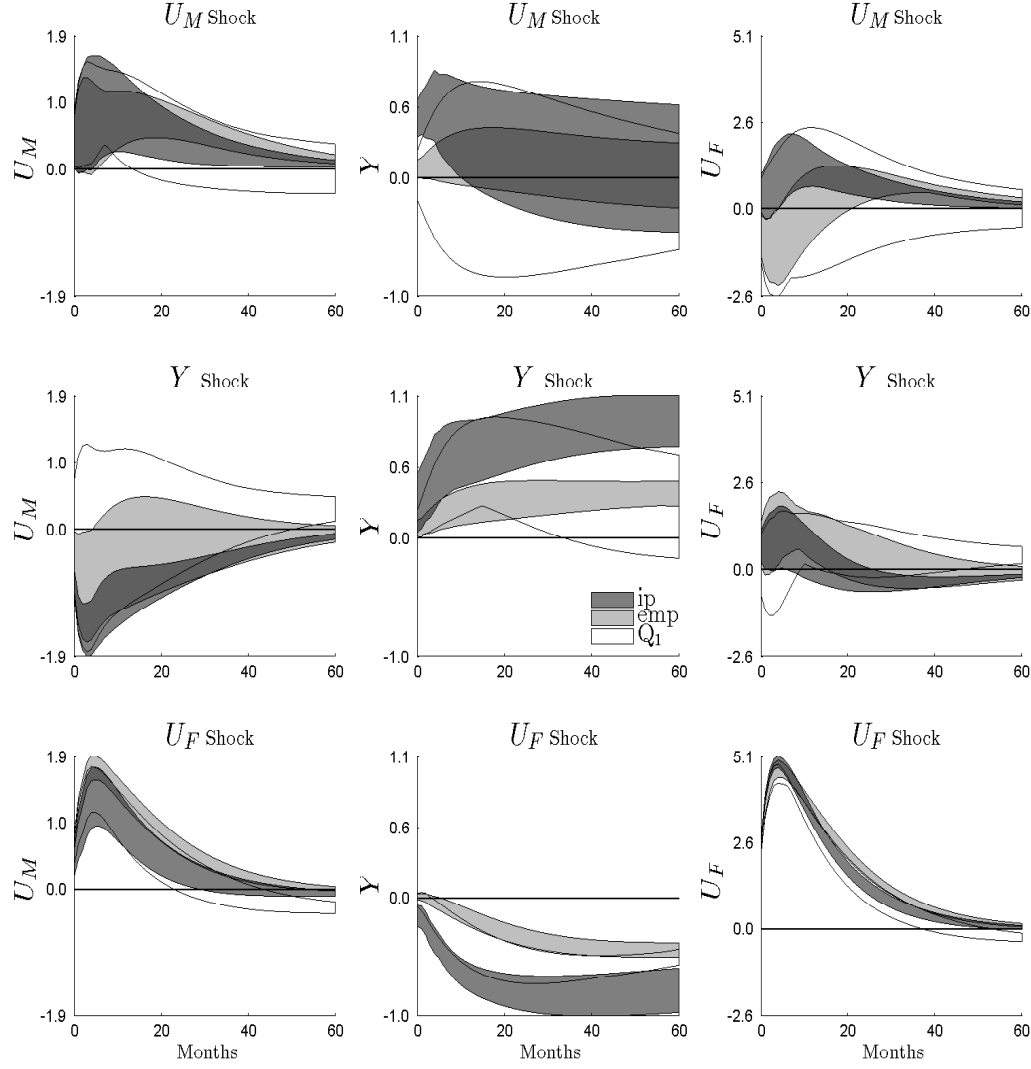
**Figure 6: IRFs of SVAR  $(U_M, ip, U_F)'$  under Alternative Bounds**



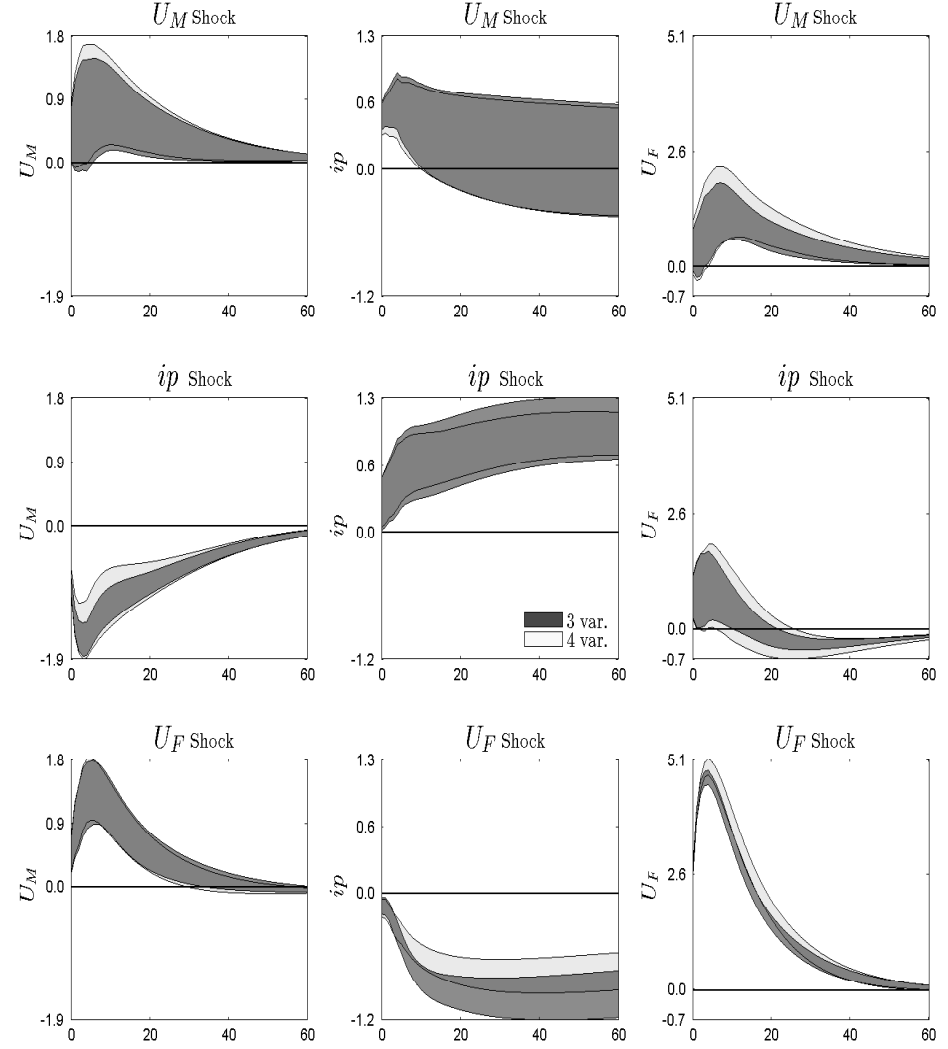
The left panel reports sets of solutions obtained when the event parameters  $(\bar{k}_1, \bar{k}_2, \bar{k}_3)$  are less restrictive while  $\bar{c}$  and  $\bar{C}$  are held fixed at their baseline values. The right panel reports sets of solutions obtained when  $\bar{c}$  and  $\bar{C}$  are less restrictive while  $(\bar{k}_1, \bar{k}_2, \bar{k}_3)$  are held fixed at their baseline values. The sample spans the period 1960:07 to 2015:04.

**Figure 7: SVAR  $(U_M, Y, U_F, r_{CRSP})'$**

**IRFs of SVAR  $(U_M, Y, U_F, r_{CRSP})'$**



**IRFs of SVAR  $(U_M, ip, U_F, r_{CRSP})'$  v.s.  $(U_M, ip, U_F)'$**



The shaded areas represent sets of solutions that satisfy the correlation and event constraints. The sample spans the period 1960:07 to 2015:04.

**Table 1: Sample Statistics**

Panel A: Correlations between Instruments and Shocks			
$SVAR$	$(U_M, ip, U_F)'$	$(U_M, emp, U_F)'$	$(U_M, Q_1, U_F)'$
$\rho(Z_{1t}(\beta), \hat{e}_{Mt})$	-0.0309	-0.0501	-0.0535
$\rho(Z_{1t}(\beta), \hat{e}_{Ft})$	-0.1853	-0.1866	-0.1809
$\rho(Z_{2t}(\beta), \hat{e}_{Ft})$	-0.1637	-0.1665	-0.1637
$\rho(Z_{1t}(\beta), \hat{e}_{Yt})$	0.0000	0.0000	0.0000
$\rho(Z_{2t}(\beta), \hat{e}_{Yt})$	0.0000	0.0000	0.0000
$\rho(Z_{2t}(\beta), \hat{e}_{Mt})$	0.0000	0.0000	0.0000
Panel B: Estimates of $\Sigma$			
$\sigma_{MM}$	0.0038	0.0030	0.0030
	[0.000, 0.007]	[0.000, 0.008]	[0.000, 0.007]
$\sigma_{YY}$	0.0035	0.0006	0.0008
	[0.000, 0.005]	[0.000, 0.002]	[0.000, 0.002]
$\sigma_{FF}$	0.0264	0.0263	0.0263
	[0.024, 0.027]	[0.023, 0.026]	[0.022, 0.027]

For the max-C solution, panel A reports the correlation between the estimated uncertainty shocks and the instruments. Panel B reports estimates of  $\Sigma$ . Lower and upper bounds across solutions that survive the correlation and event constraints are reported in square brackets. The data are monthly and span the period 1960:07 to 2015:04.



**Table 2: Variance Decomposition**

Panel A: Variance Decomposition $(U_M, Y, U_F)'$									
$s$	SVAR $(U_M, ip, U_F)'$			SVAR $(U_M, emp, U_F)'$			SVAR $(U_M, Q_1, U_F)'$		
	Fraction variation in $U_M$			Fraction variation in $U_M$			Fraction variation in $U_M$		
	$U_M$ Shock	$ip$ Shock	$U_F$ Shock	$U_M$ Shock	$emp$ Shock	$U_F$ Shock	$U_M$ Shock	$Q_1$ Shock	$U_F$ Shock
1	0.103	0.654	0.243	0.069	0.660	0.271	0.067	0.722	0.212
12	0.104	0.497	0.399	0.098	0.485	0.417	0.116	0.570	0.314
$\infty$	0.136	0.518	0.345	0.189	0.469	0.342	0.268	0.457	0.274
$s_{\max}$	0.143	0.654	0.399	0.189	0.675	0.417	0.292	0.733	0.321
	[0.01, 0.51]	[0.44, 0.79]	[0.13, 0.61]	[0.05, 0.53]	[0.02, 0.72]	[0.42, 0.70]	[0.16, 0.62]	[0.16, 0.78]	[0.23, 0.71]
	Fraction variation in $ip$			Fraction variation in $emp$			Fraction variation in $Q_1$		
	$U_M$ Shock	$ip$ Shock	$U_F$ Shock	$U_M$ Shock	$emp$ Shock	$U_F$ Shock	$U_M$ Shock	$Q_1$ Shock	$U_F$ Shock
1	0.674	0.322	0.004	0.800	0.197	0.003	0.844	0.136	0.020
12	0.321	0.453	0.226	0.553	0.311	0.136	0.521	0.223	0.256
$\infty$	0.051	0.551	0.398	0.080	0.485	0.435	0.140	0.472	0.388
$s_{\max}$	0.694	0.551	0.401	0.809	0.485	0.435	0.852	0.472	0.401
	[0.34, 0.97]	[0.18, 0.73]	[0.30, 0.61]	[0.08, 0.96]	[0.23, 0.99]	[0.40, 0.69]	[0.05, 0.97]	[0.21, 1.00]	[0.31, 0.55]
	Fraction variation in $U_F$			Fraction variation in $U_F$			Fraction variation in $U_F$		
	$U_M$ Shock	$ip$ Shock	$U_F$ Shock	$U_M$ Shock	$emp$ Shock	$U_F$ Shock	$U_M$ Shock	$Q_1$ Shock	$U_F$ Shock
1	0.002	0.033	0.965	0.018	0.033	0.948	0.008	0.035	0.957
12	0.010	0.022	0.968	0.016	0.027	0.957	0.069	0.032	0.900
$\infty$	0.024	0.030	0.946	0.069	0.030	0.900	0.170	0.048	0.781
$s_{\max}$	0.024	0.055	0.973	0.069	0.051	0.963	0.173	0.052	0.961
	[0.02, 0.19]	[0.02, 0.18]	[0.87, 0.99]	[0.05, 0.26]	[0.02, 0.23]	[0.79, 0.96]	[0.11, 0.35]	[0.04, 0.15]	[0.74, 0.97]
Panel B: Variance Decomposition for SVAR in System $(U_R, Y, U_F)'$									
$s$	SVAR $(U_R, ip, U_F)'$			SVAR $(U_R, emp, U_F)'$			SVAR $(U_R, Q_1, U_F)'$		
	Fraction variation in $U_R$			Fraction variation in $U_R$			Fraction variation in $U_R$		
	$U_R$ Shock	$ip$ Shock	$U_F$ Shock	$U_R$ Shock	$emp$ Shock	$U_F$ Shock	$U_R$ Shock	$Q_1$ Shock	$U_F$ Shock
1	0.002	0.956	0.042	0.069	0.595	0.336	0.002	0.879	0.119
12	0.005	0.755	0.240	0.033	0.440	0.527	0.008	0.699	0.293
$\infty$	0.005	0.710	0.284	0.032	0.395	0.573	0.044	0.670	0.286
$s_{\max}$	0.006	0.971	0.284	0.098	0.595	0.573	0.044	0.897	0.298
	[0.01, 0.01]	[0.97, 0.97]	[0.28, 0.28]	[0.04, 0.47]	[0.19, 0.69]	[0.55, 0.65]	[0.04, 0.80]	[0.23, 0.92]	[0.20, 0.30]
	Fraction variation in $ip$			Fraction variation in $emp$			Fraction variation in $Q_1$		
	$U_R$ Shock	$ip$ Shock	$U_F$ Shock	$U_R$ Shock	$emp$ Shock	$U_F$ Shock	$U_R$ Shock	$Q_1$ Shock	$U_F$ Shock
1	0.846	0.048	0.105	0.709	0.287	0.004	0.919	0.009	0.073
12	0.490	0.100	0.410	0.662	0.219	0.119	0.637	0.004	0.359
$\infty$	0.194	0.133	0.673	0.399	0.114	0.487	0.395	0.004	0.602
$s_{\max}$	0.848	0.133	0.673	0.713	0.298	0.487	0.927	0.011	0.602
	[0.85, 0.85]	[0.13, 0.13]	[0.67, 0.67]	[0.06, 0.87]	[0.13, 0.98]	[0.34, 0.52]	[0.12, 0.93]	[0.01, 0.82]	[0.60, 0.64]
	Fraction variation in $U_F$			Fraction variation in $U_F$			Fraction variation in $U_F$		
	$U_R$ Shock	$ip$ Shock	$U_F$ Shock	$U_R$ Shock	$emp$ Shock	$U_F$ Shock	$U_R$ Shock	$Q_1$ Shock	$U_F$ Shock
1	0.097	0.001	0.902	0.082	0.112	0.806	0.058	0.032	0.910
12	0.129	0.003	0.868	0.041	0.113	0.845	0.119	0.028	0.853
$\infty$	0.129	0.025	0.846	0.039	0.093	0.868	0.239	0.046	0.715
$s_{\max}$	0.132	0.025	0.909	0.083	0.138	0.868	0.239	0.047	0.914
	[0.13, 0.13]	[0.03, 0.03]	[0.91, 0.91]	[0.06, 0.37]	[0.01, 0.22]	[0.77, 0.88]	[0.13, 0.29]	[0.03, 0.15]	[0.91, 0.94]

For the max-C solution, each panel shows the fraction of  $s$ -step-ahead forecast-error variance of the variable given in the panel title that is explained by the shock named in the column heading. The row denoted “ $s = s_{\max}$ ” reports the maximum fraction (across all VAR forecast horizons  $m$ ) of forecast error variance explained by the shock listed in the column heading. Lower and upper bounds for the  $s = s_{\max}$  horizon across solutions that survive the correlation and event constraints are reported in square brackets. The data are monthly and span the period 1960:07 to 2015:04.

## Appendix for Online Publication

This appendix contains additional information for “Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response?”

### Closed-Form Solution for $\mathbf{B}$ when $Z$ is observed

**Lemma 1** *There exists a unique solution to the system (7) if  $\mathbb{E}[e_{Ft}Z_2] \neq 0$  and  $\mathbb{E}[e_{Mt}Z_1] \neq 0$ .*

**Proof.** To facilitate the presentation throughout the proof, let

$$\begin{aligned}\boldsymbol{\eta}_t &= \mathbf{B}\mathbf{e}_t \\ \mathbf{B} &= \begin{bmatrix} \mathbf{B}_M & \mathbf{B}_Y & \mathbf{B}_F \\ 3 \times 1 & 3 \times 1 & 3 \times 1 \end{bmatrix} \\ \boldsymbol{\Omega} &= \mathbb{E}(\boldsymbol{\eta}_t \boldsymbol{\eta}_t').\end{aligned}$$

Let  $\phi_{1F} = c_{1F}\sigma_{Z1}$ ,  $\phi_{2F} = c_{2F}\sigma_{Z2}$ ,  $\phi_{1M} = c_{1M}\sigma_{Z1}$ . We have two external instruments  $(Z_1, Z_2)$  satisfying

$$\begin{aligned}\mathbb{E}[e_{Ft}Z_1] &\equiv \phi_{1F} \neq 0, \mathbb{E}[e_{Mt}Z_1] \equiv \phi_{1M} \neq 0 \text{ and } \mathbb{E}[e_{Yt}Z_1] = 0 \\ \mathbb{E}[e_{Ft}Z_2] &\equiv \phi_{2F} \neq 0 \text{ and } \mathbb{E}[e_{Mt}Z_2] = \mathbb{E}[e_{Yt}Z_2] = 0\end{aligned}$$

Then

$$\mathbb{E}[\boldsymbol{\eta}_t Z_2] = \mathbb{E}[\mathbf{B}\mathbf{e}_t Z_2] = \mathbf{B} \begin{bmatrix} 0 \\ 0 \\ \phi_{2F} \end{bmatrix} = \phi_{2F} \mathbf{B}_F \quad (\text{A.1})$$

Thus  $\mathbf{B}_F$  exists if  $\phi_{2F} \neq 0$ . Observe that, since

$$\boldsymbol{\Omega} = \mathbb{E}[\boldsymbol{\eta}_t \boldsymbol{\eta}_t'] = \mathbf{B}\mathbf{B}'$$

we have

$$\mathbf{B}'\boldsymbol{\Omega}^{-1}\mathbf{B} = \mathbf{I}$$

hence,  $\forall i, j = M, Y, F$

$$\mathbf{B}_j' \boldsymbol{\Omega}^{-1/2} \boldsymbol{\Omega}^{-1/2} \mathbf{B}_i = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

Therefore,

$$\mathbb{E}[\boldsymbol{\eta}_t Z_2]' \boldsymbol{\Omega}^{-1} \mathbb{E}[\boldsymbol{\eta}_t Z_2] = (\phi_{2F} \mathbf{B}_F)' \boldsymbol{\Omega}^{-\frac{1}{2}} \boldsymbol{\Omega}^{-\frac{1}{2}} (\phi_{2F} \mathbf{B}_F) = \phi_{2F}^2$$

This implies that the scale  $\phi_{2F}$  is identified up to a sign by

$$\phi_{2F} = \pm \sqrt{\mathbb{E}[\boldsymbol{\eta}_t Z_2] \boldsymbol{\Omega}^{-1} \mathbb{E}[\boldsymbol{\eta}_t Z_2]}. \quad (\text{A.2})$$

Next,

$$\mathbb{E}[\boldsymbol{\eta}_t Z_1] = \mathbb{E}[\mathbf{B}\mathbf{e}_t Z_1] = \mathbf{B} \begin{bmatrix} \phi_{1M} \\ 0 \\ \phi_{1F} \end{bmatrix} = \phi_{1M} \mathbf{B}_M + \phi_{1F} \mathbf{B}_F$$

But note that

$$\begin{aligned}\mathbb{E}[\boldsymbol{\eta}_t Z_2] \boldsymbol{\Omega}^{-1} \mathbb{E}[\boldsymbol{\eta}_t Z_1] &= \phi_{2F} \mathbf{B}_F' \boldsymbol{\Omega}^{-1} (\phi_{1M} \mathbf{B}_M + \phi_{1F} \mathbf{B}_F) \\ &= \phi_{2F} \mathbf{B}_F' (\mathbf{B}\mathbf{B}')^{-1} (\phi_{1M} \mathbf{B}_M + \phi_{1F} \mathbf{B}_F) \\ &= \phi_{2F} \phi_{1F}\end{aligned}$$

This implies that  $\phi_{1F}$  is identified as

$$\phi_{1F} = \frac{\mathbb{E}[\boldsymbol{\eta}_t Z_2] \boldsymbol{\Omega}^{-1} \mathbb{E}[\boldsymbol{\eta}_t Z_1]}{\phi_{2F}}$$

which in turn implies

$$\phi_{1M} \mathbf{B}_M = \mathbb{E}[\boldsymbol{\eta}_t Z_1] - \frac{\mathbb{E}[\boldsymbol{\eta}_t Z_2]}{\phi_{2F}} c_{1F}. \quad (\text{A.3})$$

Thus solution to  $\mathbf{B}_M$  exists if  $\phi_{1M} \neq 0$ . Furthermore, note that

$$\begin{aligned} & \left( \mathbb{E}[\boldsymbol{\eta}_t Z_1] - \frac{\mathbb{E}[\boldsymbol{\eta}_t Z_2]}{\phi_{2F}^2} \phi_{1F} \right)' \boldsymbol{\Omega}^{-1} \left( \mathbb{E}[\boldsymbol{\eta}_t Z_1] - \frac{\mathbb{E}[\boldsymbol{\eta}_t Z_2]}{\phi_{2F}^2} c_{1F} \right) \\ &= \boldsymbol{\Omega}^{-\frac{1}{2}} \mathbf{B}_M \phi_{1M}^2 \mathbf{B}_M' \boldsymbol{\Omega}^{-\frac{1}{2}} = \phi_{1M}^2 \end{aligned}$$

This implies that the parameter  $\phi_{1M}$  is identified up to a sign as

$$\phi_{1M}^2 = \left( \mathbb{E}[\boldsymbol{\eta}_t Z_1] - \frac{\mathbb{E}[\boldsymbol{\eta}_t Z_2]}{\phi_{2F}^2} c_{1F} \right)' \boldsymbol{\Omega}^{-1} \left( \mathbb{E}[\boldsymbol{\eta}_t Z_1] - \frac{\mathbb{E}[\boldsymbol{\eta}_t Z_2]}{\phi_{2F}^2} \phi_{1F} \right). \quad (\text{A.4})$$

It only remains to identify  $\mathbf{B}_Y$ .  $\mathbf{B}_Y$  must satisfy

$$\begin{aligned} \mathbf{B}_Y' \boldsymbol{\Omega}^{-1/2} \boldsymbol{\Omega}^{-1/2} \mathbf{B}_Y &= 1 \\ \mathbf{B}_Y' \boldsymbol{\Omega}^{-1/2} \boldsymbol{\Omega}^{-1/2} \mathbf{B}_M &= 0 \\ \mathbf{B}_Y' \boldsymbol{\Omega}^{-1/2} \boldsymbol{\Omega}^{-1/2} \mathbf{B}_F &= 0 \end{aligned} \quad (\text{A.5})$$

$\mathbf{B}_Y$  can be solved analytically using (A.5) provided that  $\mathbf{B}_F$  and  $\mathbf{B}_Y$  are identified. In addition, since the equation (A.5) is quadratic in  $\mathbf{B}_Y$ ,  $\mathbf{B}_Y$  is unique up to sign. It follows that there exists a  $\tau$  such that

$$\mathbf{B}_Y = \tau \tilde{\mathbf{B}}_Y \quad (\text{A.6})$$

where  $\tilde{\mathbf{B}}_Y$  is unique conditional on  $\phi_{2F}$  and  $\phi_{1M}$ , but the scalar  $\tau$  is unique up to sign.

This shows that the solution to the system (7) exists and is unique up to sign if  $\phi_{2F} \neq 0$ ,  $\phi_{1M} \neq 0$ . Combined with unit effect normalization (5) and the restriction on the admissible parameter space (4),  $\mathbf{B}$  can be uniquely identified. The unit effect normalization implies

$$\begin{aligned} \begin{pmatrix} B_{MM} & B_{MY} & B_{MF} \\ B_{YM} & B_{YY} & B_{YF} \\ B_{FM} & B_{FY} & B_{FF} \end{pmatrix} &= \begin{pmatrix} 1 & H_{MY} & H_{MF} \\ H_{YM} & 1 & H_{YF} \\ H_{FM} & H_{FY} & 1 \end{pmatrix} \begin{pmatrix} \sigma_{MM} & 0 & 0 \\ 0 & \sigma_{YY} & 0 \\ 0 & 0 & \sigma_{FF} \end{pmatrix} \\ &= \begin{pmatrix} \sigma_{MM} & H_{MY}\sigma_{YY} & H_{MF}\sigma_{FF} \\ H_{YM}\sigma_{MM} & \sigma_{YY} & H_{YF}\sigma_{FF} \\ H_{FM}\sigma_{MM} & H_{FY}\sigma_{YY} & \sigma_{FF} \end{pmatrix} \end{aligned}$$

Combined with the restriction  $\sigma_{jj} > 0$  for all  $j = M, Y, F$ , implies  $B_{jj} > 0$  for all  $j = M, Y, F$ . From equation (A.1),  $B_{FF} > 0$  pins down the sign of  $\phi_{2F}$  conditional  $\mathbf{Z}_t$ . Since the sign of  $\phi_{2F}$  is pinned down, the signs of  $B_{MF}$  and  $B_{YF}$  are also pinned down by the same restriction. From equation (A.3),  $B_{MM} > 0$  pins down the sign of  $\phi_{1M}$  conditional  $\mathbf{Z}_t$  and therefore the signs of  $B_{YM}$  and  $B_{FM}$  are pinned down by the same restriction. It only remains to show the uniqueness of  $\mathbf{B}_Y$ . Provided that  $\mathbf{B}_F$  and  $\mathbf{B}_Y$  are identified and given the closed-form solution (A.5) that is quadratic in  $\mathbf{B}_Y$ , then  $B_{YY} > 0$  pins down the sign of  $\tau$  conditional  $\mathbf{Z}_t$  and hence the sign of  $B_{MY}$  and  $B_{FY}$  are also pinned down by the same restriction. ■

The system of equations defining  $\mathbf{B}$  is

$$0 = \mathbb{E}[\mathbf{g}_1(\mathbf{m}_{1t}; \boldsymbol{\beta}_1)] \equiv \bar{\mathbf{g}}_1.$$