Identifying Shocks via Time-Varying Volatility

Daniel J. Lewis

Federal Reserve Bank of New York*

October 16, 2018

Abstract

An $n-$variable structural vector auto-regression (SVAR) can be identified (up to shock order) from the evolution of the residual covariance across time if the structural shocks exhibit heteroskedasticity (Rigobon (2003), Sentana & Fiorentini (2001)). However, the path of residual covariances is available only under specific parametric assumptions on the variance process. I propose a new identification argument that identifies the SVAR up to shock orderings using the autocovariance structure of second moments of the residuals, implied by an arbitrary stochastic process for the shock variances. These higher moments are available without parametric assumptions like those required by existing approaches. I offer intuitive criteria to select among shock orderings; this selection does not impact inference asymptotically. The identification scheme performs well in simulations. I apply it to the debate on fiscal multipliers. I obtain estimates lower than those of Blanchard & Perotti (2002) and Mertens & Ravn (2014), but in line with more recent studies.

Keywords: identification, impulse response function, structural shocks, SVAR, fiscal multiplier, time-varying volatility, heteroskedasticity.

JEL codes: C32, C58, E20, E62, H30.

---

*The views expressed in this paper are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. Helpful comments provided by Jim Stock, Gabriel Chodorow-Reich, Neil Shephard, Mikkel Plagborg-Møller, Lutz Kilian, Jón Steinsson, Giri Parameswaran, Domenico Giannone, Richard Crump, Andrew Patton, and participants at various seminars.

1 Introduction

The central challenge of structural vector autoregression (SVAR) analysis is to identify underlying structural shocks from observable VAR innovations (one-step ahead reduced-form forecast errors). For example, an innovation to tax revenues could represent either a true tax shock or the effect of automatic stabilizers as a response to changing macroeconomic conditions. Policy analysis centers on the relationships between structural shocks and observables. In a SVAR, the reduced-form innovations, η_t, are expressed as a linear combination of the underlying shocks, ε_t: $\eta_t = H\varepsilon_t$ for some contemporaneous response matrix H. Up to second moments, these equations have a multiplicity of solutions for H; economic assumptions are generally needed for identification. The majority of approaches use "internal instruments", restricting elements of H to identify the remainder. These restrictions can be short-run exclusions (Sims (1980)), long-run exclusions (Blanchard & Quah (1986)), on signs (Uhlig (2005)), or calibrated parameters (Blanchard & Perotti (2002)). More recently, "external instruments" have been proposed as an alternative, as in Mertens & Ravn (2013). However, these assumptions are frequently controversial.

A smaller literature offers identification based on statistical properties of the innovations. Sentana & Fiorentini (2001) and Rigobon (2003) share the important insight that if the variances of structural shocks change over time, shocks can be identified from the reduced-form covariances at different points in time. However, this path of reduced form covariances is available only under specific parametric models. The method of Rigobon fits discrete variance regimes to the data, either based on external information or estimation. Sentana & Fiorentini (2001) use the full path of covariances, feasible only under models like GARCH. Generalizations have been made to Markov switching (Lanne, Lütkepohl, & Maciejowska (2010)) and smooth transitions between regimes (Lütkepohl & Netšunajev (2017)). All of these approaches rely on knowledge of the path of variances over time or parametric features allowing that path to be consistently estimated, which has so far limited researchers to choose one of the few models that can be accommodated. There is compelling evidence of time-varying volatility in US macroeconomic aggregates, as documented by Stock & Watson (2002), Blanchard & Simon (2001), and Jurado, Ludvigson, & Ng (2015), so identification based on heteroskedasticity has the potential to be very useful in practice.

I present a new identification argument based on heteroskedasticity that does not refer to the variance path, and thus need not make use of a particular parametric model. If time-varying volatility is present, in any (unspecified) form, identification follows from the autocovariance of the volatility process. Since shocks are assumed to be uncorrelated over time, the autocovariance of squared residuals picks up only dynamics of the volatility process. This autocovariance furnishes equations that identify the response matrix and the structural shocks (up to an ordering), under very general conditions. In a simple model, the use of the autocovariance for identification can be motivated as an instrumental variables problem. The argument is similar in spirit to the non-Gaussian identification of e.g. Gouriéroux & Monfort (2015, 2017) and Hyvärinen, Zhang, Shimizu, & Hoyer (2010), which uses different higher moments, but requires assumptions on the shocks that rule out common forms of heteroskedasticity.

Identification based on time-varying volatility (TVV-ID) should be seen as an argument establishing identification under general conditions. Indeed, it separately establishes identification via a novel channel for the models that have previously been shown to offer identification via heteroskedasticity (e.g. GARCH of Sentana & Fiorentini (2001) and regimes of Rigobon (2003)). More importantly, it gives researchers the freedom to develop new alternative models and procedures in contexts exhibiting time-varying volatility, without having to stop to establish identification from scratch. As opposed to identification via heteroskedasticity being a model-dependent argument, TVV-ID progresses towards a model-free argument, which researchers can apply in ways that best suit their data.

This means that any estimator that fits an autocovariance to the variance process can implement TVV-ID. The most natural candidate is GMM, which needs no parametric assumptions. However, a researcher can also use a (quasi-) likelihood using any model that implies such an autocovariance. I compare a variety of approaches considered in the literature with some newly-admissible based on TVV-ID. I find that an estimator based on an AR(1) stochastic volatility (SV) model performs best across many DGPs. I also argue that the Rigobon approach, when regimes must be estimated (the most comparably-agnostic identification approach to TVV-ID) is susceptible to bias.

Identification via heteroskedasticity has been widely adopted in practice. Its use has spread from macrofinance to fields including public finance, growth, trade, political economy, agriculture, energy, education, marketing, and even fertility. This proliferation illustrates that there is potential value in freeing applied researchers from the strict parametric models they have been required to use, and understanding any limitations of such models. The full flexibility offered by TVV-ID also shows that macro models including time-varying volatility are often estimated without realizing and exploiting its implications for identification. For example, Primiceri (2005) assumes a triangular H matrix when his volatility model means doing so unnecessarily restrictive.

As an empirical application, I use TVV-ID to estimate fiscal multipliers and test previous identifying assumptions from the literature.[1] The multipliers I estimate are lower than those of Blanchard & Perotti (2002) or the comparative study of Mertens & Ravn (2014). I show that the narrative tax shocks often used for identification may not pass standard tests for validity. I reject the key parameter, the elasticity of tax revenues to output, obtained by both and obtain a value, 1.58, in line with Follette & Lutz's (2010) estimate based on institutional data. My multipliers accord with recent estimates of Caldara & Kamps (2017) and Ramey & Zubairy (2018).

The remainder of this paper proceeds as follows. Section 2 describes the identification problem in detail and presents the theoretical results. Section 3 addresses the interpretation of results from TVV-ID. Section 4 compares of implementations of TVV-ID and other identification schemes in simulation. The empirical application follows in Section 5. Section 6 concludes.

2 Identification theory

In the canonical SVAR setting, a vector of innovations, η_t, is composed of unobserved structural shocks, ε_t, via a response matrix, H. More broadly, this represents a decomposition problem. η_t is $n \times 1$, obtained from a reduced-form model, or directly observed. For example, a structural vector auto-regression (SVAR) based on data Y_t

[1] I have considered numerous other empirical applications. To summarize key results, I find that the recursive structure of Bernanke, Boivin, & Eliasz (2005) can be rejected, causing the price puzzle to return, and promoting other surprising behaviour at the contemporaneous horizon; the recursive structure of Kilian (2009) summarizes the data well (assumed zeros correspond to precisely estimated zeros); in Kilian & Park (2009) the zeroes assumed in the asset column of the contemporaneous response matrix are at odds with point estimates, but cannot be rejected; the assumptions of Blanchard & Quah (1989) are borne out strongly by TVV-ID; the exogeneity of uncertainty assumed in Bloom (2009) can be rejected, and the shapes of key responses to uncertainty shocks change somewhat.

would yield $A(L)Y_t = \eta_t$. Similarly, ε_t is $n \times 1$, so H is $n \times n$. Thus,

$$\eta_t = H\varepsilon_t, t = 1, \ldots, T, \tag{1}$$

leaving H and, equivalently, ε_t, to be identified. Alternatively, (1) describes a factor model, for example. I present a simple example under special assumptions to outline the identification problem and how heteroskedasticity may solve it. I then derive a representation of higher moments of the reduced-form innovations to serve as identifying equations. The following section establishes conditions under which these equations have a unique solution. I discuss possibly restrictive assumptions and the relation to previous approaches.

2.1 Intuition for the use of heteroskedasticity

Before the impact of heteroskedasticity can be illustrated, assumptions underlying Equation (1), are required.

Assumption 0. (temporary) *For all $t = 1, 2, \ldots, T$,*

1. $E_t[\varepsilon_t\varepsilon_t' \mid \sigma_t] = diag(\sigma_t^2) \equiv \Sigma_t$ *(σ_t is the conditional variance of the shocks),*
2. *σ_t is a fourth-order stationary strictly positive stochastic process,*
3. $E[\Sigma_t] = \Sigma_\varepsilon$,
4. *Shocks satisfy conditional mean independence, $E[\varepsilon_{it} \mid \varepsilon_{-is}] = 0$ for all i, all $t, s = 1, 2, \ldots T$,*
5. *H is time-invariant, invertible, with a unit diagonal normalization.*

The fourth point substitutes conditional mean independence for the usual slightly weaker uncorrelated shocks assumption. While the variance of shocks may change, fixing H means that the economic impact of a unit shock remains the same. It is natural to seek to identify H from the overall covariance of η_t, $E[\eta_t\eta_t'] = \Sigma_\eta$. However, it is well-known that these equations can only identify H up to an orthogonal rotation, Φ ($\Phi\Phi' = I$).[2]

[2]Observe $\Sigma_\eta = H\Sigma_\varepsilon H' = (H\Phi)(\Phi'\Sigma_\varepsilon\Phi)(H\Phi)' = H^*\Sigma_\varepsilon^* H^{*'}$, where $H^* = H\Phi D_{H,\Phi}$ and $\Sigma_\varepsilon^* = D_{H,\Phi}^{-1}\Phi'\Sigma_\varepsilon\Phi D_{H,\Phi}^{-1}$, with $D_{H,\Phi}$ the matrix that unit-normalizes the diagonal of $H\Phi$. This means that

Variation in Σ_t may allow the researcher to overcome this. Consider a simple two-variable example, where one structural variance is time-varying and the other is fixed. This admits the simplest form of the Rigobon approach, which yields closed form solutions for H (see e.g. Nakamura & Steinsson (2018)). Without loss of generality, assume σ_{2t}^2 changes and $\sigma_{1t}^2 \equiv \sigma_1^2$, constant. Denote

$$\sigma_t^2 = \begin{bmatrix} \sigma_1^2 \\ \sigma_{2t}^2 \end{bmatrix}, \; H = \begin{bmatrix} 1 & H_{12} \\ H_{21} & 1 \end{bmatrix}.$$

The conditional variances of the reduced-form innovations are given by $E_t [\eta_t \eta_t' \mid \sigma_t] = H \Sigma_t H'$. Given two subsamples, A, B, containing the sets of time points T_A, T_B, it is shown by Nakamura & Steinsson (2018) (and in the Supplement) that

$$\frac{E_{T_A}[\eta_{1t}\eta_{2t}] - E_{T_B}[\eta_{1t}\eta_{2t}]}{E_{T_A}[\eta_{2t}^2] - E_{T_B}[\eta_{2t}^2]} = \frac{H_{12}\Delta\left(\sigma_{2t}^2\right)}{\Delta\left(\sigma_{2t}^2\right)} = H_{12}. \tag{2}$$

where the $\Delta(\cdot)$ operator represents the difference in expectation of the argument between subsamples T_A, T_B. Assuming that $\Delta\left(\sigma_{2t}^2\right) \neq 0$, H_{12} can thus be identified in closed form. σ_{2t} need only have finite second moments for all $t \in T_A, T_B$. While the Rigobon identification scheme is motivated by a regime-based process, identification holds even when misspecified, provided $\Delta\left(\sigma_{2t}^2\right) \neq 0$, and σ_1 is indeed fixed. If there are regimes, they need not be known or correctly specified, as noted in Rigobon (2003). However, if the value of the σ_{2t} process is instead constant, $\Delta\left(\sigma_{2t}^2\right)$ would be zero in population, and identification fails.[3]

Rigobon's approach provides moment conditions based on means of the variance process, which can yield identification for many processes, but arguments are possible using other moments. Across periods, there is motivation for an instrumental variables

the pairs (H, Σ_ε) and $(H^*, \Sigma_\varepsilon^*)$ are observationally equivalent. Alternatively, note that due to the symmetry of Σ_η, it offers $n(n+1)/2$ equations, but there are n^2 unknowns. This is the fundamental identification problem posed by the SVAR methodology and indeed many related models (e.g. factor models).

[3]In this case, if regimes are instead estimated from the values of η_t, the resulting estimates of $\Delta\left(\sigma_{2t}^2\right)$ are not zero in population since regime selection is driven by realized shock values, but this source of variation results in bias, as discussed in Section 4.

approach. Noting

$$\begin{aligned}\eta_{2t}\eta_{1t} &= H_{21}\varepsilon_{1t}^2 + H_{12}\varepsilon_{2t}^2 + \varepsilon_{1t}\varepsilon_{2t} + H_{12}H_{21}\varepsilon_{1t}\varepsilon_{2t},\\ \eta_{2t}^2 &= H_{21}^2\varepsilon_{1t}^2 + 2H_{21}\varepsilon_{1t}\varepsilon_{2t} + \varepsilon_{2t}^2,\end{aligned}$$

it is clear that H_{12} would be identified from the ratio of the $H_{12}\varepsilon_{2t}^2$ and ε_{2t}^2 terms. This is not possible as only the values of η_t are observed, and not their separate components. However, a lagged value of η_{2t}^2 can be an instrument for ε_{2t}^2. Note

$$cov\left(\eta_{2t}\eta_{1t}, \eta_{2(t-p)}^2\right) cov\left(\eta_{2t}^2, \eta_{2(t-p)}^2\right) = H_{12},\ cov\left(\varepsilon_{2t}^2, \varepsilon_{2(t-p)}^2\right), = cov\left(\varepsilon_{2t}^2, \varepsilon_{2(t-p)}^2\right),$$

by Assumption 0.4 and the fact that σ_1 is fixed. H_{12} is then identified in closed form:

$$\frac{cov\left(\eta_{2t}\eta_{1t}, \eta_{2(t-p)}^2\right)}{cov\left(\eta_{2t}^2, \eta_{2(t-p)}^2\right)} = \frac{H_{12}cov\left(\varepsilon_{2t}^2, \varepsilon_{2(t-p)}^2\right)}{cov\left(\varepsilon_{2t}^2, \varepsilon_{2(t-p)}^2\right)} = H_{12}. \tag{3}$$

This is the familiar IV estimator, where the dependent variable is $\eta_{2t}\eta_{1t}$, the endogenous regressor is η_{2t}^2, and the instrument is $\eta_{2(t-p)}^2$. This works because the previous value $\eta_{2(t-p)}^2$ is uncorrelated with all period t terms except those containing ε_{2t}^2. The argument applies for any lag, p. σ_{2t} is assumed to be fourth-order stationary (for expositional simplicity) and $E\left[\varepsilon_{2t}^4\right] < \infty$. Identification holds provided

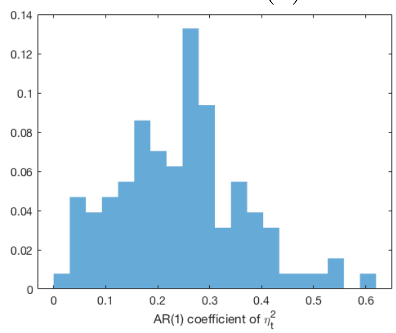
$$cov\left(\varepsilon_{2t}^2, \varepsilon_{2(t-p)}^2\right) \neq 0$$

for some p.

This requirement that the p^{th} autocovariance of η_{2t}^2 is non-zero is satisfied by a variety of processes for σ_{2t}^2. If the true process is stochastic regime-switching, it follows from the non-zero autocovariance around break dates. In a SV model, it holds if the AR coefficient is non-zero. In a GARCH model at least one of the auto-regressive parameters must be non-zero. This is the crux of TVV-ID: given the structure of the autocovariance of $\eta_t\eta_t'$, comparing elements of the autocovariance (in this simple case, via a ratio) identifies the columns of H.

This flexibility of identification – independent of specification – is not shared by the existing approaches. I have made no assumptions about whether the heteroskedasticity is conditional or unconditional (either can imply a suitable autocovariance).

Figure 1: Distribution of AR(1) coefficients of η_t^2



Time series η_t are obtained as reduced-form innovations from AR(12) processes fitted to each of McCracken & Ng's 128 FRED-MD monthly time series. The figure displays the distribution of the implied AR(1) coefficients of η_t^2.

Section 4 demonstrates that the performance of both the Sentana & Fiorentini and Rigobon approaches suffer under misspecification of the parametric model. In contrast, in this example, I have required that the stochastic process is stationary and exhibits some degree of persistence.

Empirically there is strong evidence of such persistence, as discussed in Jurado et al (2015), for example. Figure 1 displays AR(1) parameters of η_t^2, where η_t are residuals of AR(12) models fitted to each series McCracken & Ng's FRED-MD database in turn. I reject the null hypothesis of zero autocovariance at the 1% level for 96 of the 128 series, 5% for 98, and 10% for 101. A Ljung-Box test, as in Lanne & Saikkonen (2007), for $\ln T$ autocovariances rejects homoskedasticity at the 1% level for 100 of the series and the 5% level for 103. The identifying condition is frequently satisfied empirically.

Multiple autocovariances can easily be combined; each yields moments of the form

$$cov\left(\eta_{2t}\eta_{1t}, \eta_{2t-p}^2\right) - H_{12}cov\left(\eta_{2t}^2, \eta_{2t-p}^2\right) = 0$$

which can be stacked to yield an overidentified GMM problem. Alternatively, it might be natural to assume that the variances follow some loose parametric form, like an AR(1), and let this imply the whole series of autocovariances.

2.2 Identification via time-varying volatility

To this point, I have made strong assumptions to assist intuition. I now relax them and develop TVV-ID in its general form. Again, let

$$\eta_t = H\varepsilon_t,\ t = 1, 2, \dots T.$$

Write $\mathcal{F}_{t-1}$ to denote $\varepsilon_1, \dots \varepsilon_{t-1}$ and $\sigma_1, \dots \sigma_{t-1}$. I replace Assumption 0 with Assumption A:

Assumption A. For every $t = 1, 2, \dots, T$,

1. $E_t(\varepsilon_t \mid \sigma_t, \mathcal{F}_{t-1}) = 0$ *and* $Var_t(\varepsilon_t \mid \sigma_t, \mathcal{F}_{t-1}) = \Sigma_t$,
2. $\Sigma_t = diag(\sigma_t^2), \sigma_t^2 = \sigma_t \odot \sigma_t$,
3. $E_t[\sigma_t^2] < \infty$.

In addition, I make a preliminary assumption on H:
Assumption B. H *is time-invariant.*

By explicitly conditioning on σ_t, these assumptions cover both SV and autoregressive conditional heteroskedasticity-type (ARCH) models (where σ_t is a function of $\varepsilon_1, \dots \varepsilon_{t-1}$), amongst others , including unconditional heteroskedasticity.

Decomposition

To obtain moments in terms of just H and the underlying volatility process, I work with a transformation of η_t, (ζ_t, see below), as my basic data. I begin by writing the decomposition,

$$\eta_t\eta_t' = H\Sigma_t H' + V_t,\ V_t = H\left(\varepsilon_t\varepsilon_t' - \Sigma_t\right)H',$$

where σ_t^2 is unknown. Define L to be an elimination matrix, and G a selection matrix (of ones and zeros), see e.g. Magnus & Neudecker (1980).[4] Then

$$\begin{aligned}\zeta_t &= vech(\eta_t\eta_t') = vech(H\Sigma_t H') + vech(V_t) \\ &= L(H \otimes H)\,vec(\Sigma_t) + v_t,\ v_t = vech(V_t) \quad (4)\\ &= L(H \otimes H)\,G\sigma_t^2 + v_t, \quad (5)\end{aligned}$$

[4]This means $vech(A) = Lvec(A)$ and $vec(ADA') = (A \otimes A)\,Gd$ where $d = diag(D)$.

The simplification from (4) to (5) in the first term is surprising and follows due to the diagonality of Σ_t using A.2. This feature plays a key role in properties established later. From the definition of V_t, A.1, A.3, and B, $E_t[V_t \mid \sigma_t, \mathcal{F}_{t-1}] = 0$, so $E_t[v_t \mid \sigma_t, \mathcal{F}_{t-1}] = 0$ and

$$E_t[\zeta_t \mid \sigma_t, \mathcal{F}_{t-1}] = L(H \otimes H) G\sigma_t^2.$$

This provides a signal-noise interpretation for the decomposition of the outer product $\eta_t \eta_t'$. It follows from A.3 that I can integrate over Σ_t to obtain $E_t[v_t \mid \mathcal{F}_{t-1}] = 0$ and similarly that $E_t[|v_t|] < \infty$. Therefore v_t is a martingale difference sequence.

Properties of ζ_t

Coupled with the decomposition derived above, Assumption C expands on A.3 to establishuseful properties of $\zeta_t = vech(\eta_t \eta_t')$.

Assumption C. For every t,

1. $Var_t(\sigma_t^2) < \infty$,
2. $Var_t(\varepsilon_t \varepsilon_t') < \infty$.

Using these additional assumptions, the autocovariance of ζ_t has a convenient form:

Proposition 1. *Under Assumptions A.1-2, B, & C,*

$$Cov_{t,s}(\zeta_t, \zeta_s) = L(H \otimes H) GM_{t,s}(H \otimes H)' L', \; t > s \tag{6}$$

where

$$M_{t,s} = E_{t,s}\left[\sigma_t^2 \sigma_s^{2'}\right] G' + E_{t,s}\left[\sigma_t^2 vec(\varepsilon_s \varepsilon_s' - \Sigma_s)'\right] - E_t\left[\sigma_t^2\right] E_s\left[\sigma_s^{2'}\right] G'.$$

This equation represents an "observable" quantity, $\text{Cov}_{t,s}(\zeta_t, \zeta_s)$, as a product of H and the $n \times n^2$ $M_{t,s}$ (composed of moments of the underlying variance process). If $E_{t,s}\left[\sigma_{it}^2 \left(\varepsilon_s \varepsilon_s^{'} - \Sigma_s\right)\right]$ is diagonal (any ARCH effects come from own past shocks), $M_{t,s}$ can be replaced with $\tilde{M}_{t,s}G$ where $\tilde{M}_{t,s}$ is only $n \times n$.

An autocovariance of the vectorization of $\eta_t \eta_t'$ can thus be expressed as just a product of H, a nuisance matrix, and known matrices of zeros and ones. This is remarkably parsimonious for what is a covariance of random matrices. Note that stationarity has not been assumed, merely the existence of higher moments. All of the expectations used are well-defined for an object at a particular point in time, even if the distribution might be different at another point in time. A single autocovariance provides $(n^2+n)/2 \times (n^2+n)/2$ equations in $2n^2-n$ unknowns, it remains to show that this system of equations has a unique solution.

Uniqueness

Having derived a set of equations of adequate order to identify H, it remains to show that they yield a unique solution. I strengthen the assumptions on H from Assumption B:

Assumption B'. *H is time-invariant, invertible, with a unit diagonal.*[5]

Given Assumption B', the conditions under which equation (6) yields a unique solution for H are established by Theorem 1.

Theorem 1. *Under Assumptions A.1-2, B', & C, equation (6) holds. Then H and $M_{t,s}$ are jointly uniquely determined from (6) (up to labeling of shocks) provided rank$(M_{t,s}) \geq 2$ and $M_{t,s}$ has no scalar multiple rows.*

Theorem 1 states that (under certain conditions) Equation (6) will yield a unique solution for the relative magnitudes of elements in each column of H. The solution is unique up to column order, given the unit-diagonal normalization. However, there are $n!$ column orderings. The same is true for shocks identified via non-Gaussianity (Lanne & Lütkepohl (2010), amongst others), many implementations of the Rigobon approach, or identification in finite mixture models, and is discussed in Chapter 14 of Kilian & Lütkepohl (2017). In some cases, the labeling of shocks is unnecessary (as in factor models), and identification is complete, but for policy analysis labeling is required, as discussed in Section 3.

The identification argument is based on period specific moments – an autocovariance between s, t – so stationarity has not been assumed. However, in practice, for the

[5]The unit diagonal assumption is a normalization, without loss of generality.

identifying moments to be feasible in standard settings, fourth-order stationarity of σ_t or ε_t is generally needed so that the moment (6) can be consistently estimated across the full sample. In contrast, for the Sentana & Fiorentini (2001) argument to become implementable (and consistently estimable), a functional form akin to GARCH must be assumed.

Crucially, unlike previous approaches, Theorem 1 does not require knowledge of the path of Σ_t. The conditions impose interpretable restrictions on the dimension of the process σ_t^2. Jurado et al (2015) find that there are strong idiosyncratic components in time-varying volatility that cannot be explained by common factors, suggesting these assumptions are likely to hold. The rank condition is analogous to the requirement in Rigobon identification that the two regimes do not evolve proportionally. In a SV model, $M_{t,s}$ is just the autocovariance of σ_t^2 and the rank assumption is satisfied if the stochastic process σ_t^2 has at least two linearly independent dimensions. For instance, the elements of σ_t^2 cannot all depend linearly on a single common factor and idiosyncratic i.i.d. noise. If this were the case, (6) could be rewritten as $vech\,(H\Lambda H')\,vech\,(H\Lambda H')'$ for some diagonal Λ, the square of original SVAR identification problem. However, if one variance depends *nonlinearly* on the common factor, the rank condition holds.

When $n > 2$, $M_{t,s}$ must also satisfy a scalar multiple condition. This requirement is weaker than a full-rank assumption. The most obvious violation occurs if two variances are proportional. It is best regarded as a technical assumption pertaining to a pathological case where the linear algebra arguments guaranteeing uniqueness break down. In practice, there is little reason to think this condition will be violated; rather, it is more likely to lead to a weak identification problem if nearly violated. The Supplement offers a brief survey of weak identification in non-linear models. Nevertheless, In some finance settings, see eg. Campbell, Giglio, Polk, & Turley (2017), many volatilities are assumed to move proportionally. If such assumptions are merely approximations to the truth, then weak identification could result. If they are literally true, it is helpful to understand what can still be identified, which motivates the next result.

Even if the scalar multiple condition were to fail, identification is still possible for those columns of H unaffected, as shown by Corollary 1.

Corollary 1. *Under Assumptions A.1-2, B', & C, equation (6) holds. Then $H^{(j)}$ is identified from* (6) *provided $rank\,(M_{t,s}) \geq 2$ and $M_{t,s}$ contains no rows proportional* *to row j.*

The dimensionality and scalar multiple assumptions in Theorem 1 can be relaxed further by supplementing additional equations. If, for example, the (often highly informative) mean

$$E_t\left[\eta_t\eta_t'\right] = E_t\left[\zeta_t\right] \tag{7}$$

is considered, Theorem 1 can be replaced with Theorem 2.

Theorem 2. *Under Assumptions A.1-2, B', & C, equation (6) holds. Then H is uniquely determined from (6) and (7) (up to labeling of shocks) provided* $\left[\begin{array}{cc} M_{t,s} & E_t\left[\sigma_t^2\right] \end{array}\right]$ *has rank* ≥ 2 *and no scalar multiple rows.*

Theorem 2 requires that, in order for identification to fail, a scalar multiple assumption also relate $E_t\left[\sigma_t^2\right]$ to $M_{t,s}$. Similar arguments can be made, adding in further observable moments, requiring progressively more extensive scalar multiple deficiencies for identification to break down. Corollary 1 can also be extended using the logic of Theorem 2.

Overidentification and Assumption B'

Even for $n = 2$, (6) is overidentified, with the degree increasing in n. Tests exploiting overidentification can be conducted in contrast to most other identification schemes, where strong assumptions are required to yield even a just-identified model. Meaningful modeling assumptions that such tests could highlight are invertibility and that H is stable. A growing literature considers issues surrounding invertibility or nonfundamentalness, see Chapter 17 of Kilian & Lütkepohl (2017) for a review. In short, if there is not an invertible mapping between η_t and ε_t, the model cannot be identified. In practice, invertibility must almost always be assumed unless the VAR is modified, for example to allow for MA components (the recent work of Chahrour & Jurado (2017) discusses some exceptions).

While TVV-ID focuses on the instability of the variances of structural shocks, H is assumed fixed. While this may seem inconsistent, there are several points to consider. No existing identification scheme flexibly handles time-varying H (Carreiro, Clark & Marcellino (2017) and Angelini, Bacchiocchi, Caggiano, & Fanelli (2018) do so under specific functional forms). Even simple identification based on Cholesky structure, when the true structure is Cholesky, does not identify a known moment of H if H is time-varying. Compared to other schemes that assume time-varying volatility, such as Rigobon (2003) (which already uses subsamples), TVV-ID is in a better position to consider sub-sample estimation to evaluate the stability of H over time. Allowing H to vary presents an interesting econometric problem, which is a prominent part of an ongoing research agenda. However, even if H varies, provided it does so at a slower rate than the variances, identification may still hold; H will be locally stationary over intervals over which the variances are not. Theoretical work has embraced this; for example, Barro & Liao (2017) split volatility into short-run and long-run components, with agents' behaviour driven by the slower moving component. Should a researcher remain worried about the assumption of a fixed H, tests of overidentifying restrictions remain an option. Further, Andrews (1993) develops tests for parameter instability in a GMM context, for example the sup-Wald test, the conditions for which are satisfied for a variety of time-varying volatility models.[6]

Connection to signal processing and identification via non-Gaussianity TVV-ID is closely related to the signal processing literature, where the volatilities of ε_t are signals and η_t noisy measurements. This problem appears in relation to medical devices such as electroencephalograms, (see Blanco & Mulgrew (2005)) and earthquake detection (Bharadwaj, Demanet, & Fournier (2017)). The extensive work of Hyvärinen and co-authors (see e.g. Hyvärinen, Karhunen, & Oja (2001)) considers the model as a signal extraction problem, developing variants of the Independent Components Analysis approach to exploit non-Gaussianity. This research also exploits higher moments – but central moments (or cumulants) – whereas, motivated by heteroskedasticity – I focus on moments involving past values. Identification via non-Gaussianity is growing in prominence in economics (e.g. Gouriéroux & Monfort (2015, 2017)). In principle, non-Gaussianity encompasses heteroskedasticity, as time-varying volatility makes shocks unconditionally non-Gaussian. However, identification via non-Gaussianity requires that, for $i \neq j$, shocks ε_{it} and ε_{jt} be mutually independent, not just orthogonal. This rules out dependence in higher moments, and

[6]The less-familiar assumptions needed in Andrews (1993), those of Near-Epoch Dependence (NED), can be replaced by stronger properties that hold for both GARCH and SV processes. Lindner (2009) shows that GARCH satisfies β-mixing (and thus α-mixing with exponential rate) and Davis & Mikosch (2009) show that SV models inherit the mixing properties of the log-variance process. Andrews' (1983) results show that an AR(1) variance process is $\alpha-$mixing with exponential rate. These mixing properties can be shown to imply NED; see Davidson (1994) Chapter 17 for additional background.

thus any correlation in volatility across shocks, restricting the forms of heteroskedasticity that can be accommodated.

2.3 Nesting the existing approaches

TVV-ID holds in virtually any case where previously developed identification schemes apply. While Proposition 4 of Sentana & Fiorentini (2001) shows that the presence of time-varying volatility is sufficient to identify this model, conditional on the path of variances, TVV-ID demonstrates that knowing the values the variance takes is not necessary for identification. Sentana & Fiorentini's ability to apply their result is restricted by the need to know the path of $H\Sigma_t H'$ for $t = 1, ..., T$. Generically, only the noisy $\eta_t \eta_t' = H\varepsilon_t \varepsilon_t' H'$, not the moments on which the argument rests, $H\Sigma_t H'$, are available to the researcher, or meaningfully estimable. This means that to apply the result a functional form must be assumed that allows the variance path to be recovered deterministically from a single realization of $\eta_{1:T}$, H, and the GARCH parameters. To this point, the only class proposed to do so is GARCH, by Sentana & Fiorentini, where σ_t^2 is directly recoverable from past values of σ_t^2 and ε_t^2. While misspecification is often an issue for estimation, it is concerning for *identification* to rest on functional form in such a knife-edge way. However, the GARCH model is nested by the TVV-ID identification argument, using unconditional moments, as the (stationary) GARCH process clearly implies a suitable collection of matrices M_p for autocovariance p.[7] The one exception is models with a single dimension of heteroskedasticity, when TVV-ID fails but the Sentana & Fiorentini approach could identify one column of H.

In a recent paper, Bertsche & Braun (2018) estimate a stochastic volatility model for a SVAR, motivated by the identification result of Sentana & Fiorentini (2001). However, such an argument cannot alone identify the model in a meaningful sense, as it relies on moments cannot be well-estimated (the covariance of innovations period-by-period). On the basis of TVV-ID though, their model is clearly identified from the unconditional moments in (6).

Rigobon's (2003) approach simplifies Sentana & Fiorentini's argument to use the unconditional variances of two (or more) discrete variance regimes. If regime dates can be discerned using external information or estimation, then these moments are feasible based on the data. While within a regime there is zero autocovariance in

[7] Milunovich & Yang (2013) offer a (local) identification argument for the GARCH model based on reduced-form moments, more similar to the TVV-ID approach.

the volatilities, at each transition there is non-zero autocovariance. Thus, provided regime switches occur in the data, TVV-ID also establishes identification. As above, an exception occurs if only one variance changes, in which case the Rigobon scheme works but TVV-ID does not. If switches are very infrequent, it may also be the case that identification offered by TVV-ID is weak. Conversely, I argue below in Section 4.1 that the Rigobon scheme faces its own challenges when regimes are estimated – regimes are endogenous with respect to the structural shocks.

The fact that, in general, TVV-ID nests these approaches underscores its flexibility as a general identifying argument. Indeed, it empowers researchers to specify new models or approaches, secure in the knowledge that identification will follow provided there is time-varying volatility that exhibits autocovariance. This freedom allows models to be well-tailored to the data at hand on a case-by case basis, and means that heteroskedasticity can be exploited in contexts where neither GARCH nor a regime model would be satisfactory.

3 Interpreting results

Having identified H through TVV-ID, it is frequently still necessary to label the resulting structural shocks, or, equivalently, the columns of H. Kilian and Lütkepohl (2017) discuss how there may in fact be some difficulty in interpreting these as economically meaningful shocks, given the purely statistical methods used to derive them; this step helps to develop such interpretations. I discuss how labeling can proceed, argue that it does not impact asymptotic inference on H, and highlight how transparent the impact of economic assumptions can be. I sketch how labeling might proceed based on several familiar assumptions below:

– If a recursive structure is thought a reasonable approximation to the truth, the columns of H can be ordered to minimize the norm between the relevant elements and zero. This lets the data dictate more realistic *near*-zeros instead of *assuming* sharp zeros. Similar analogs exist for Uhlig's (2005) sign restrictions, or Blanchard & Quah's (1989) long-run restrictions.

– Imposing a restriction on a column of interest labels that column ex ante (as does restricting $n-1$ columns). This approach is adopted in Lanne & Lütkepohl (2010).

- Any aspect of the dynamics of the variance processes can be used to choose between shocks. This can be unconditional or with reference to historical episodes (as is frequently done with the Rigobon approach).

- A forecast error variance decomposition can label H by supposing that within a period, the majority of unpredictable variation in a particular series is driven by a certain type of shock.

- If an external instrument exists, as described by Stock (2008), the shock that is its best predictor can be selected.

- Certain magnitudes of responses can be ruled out as implausible. This is often very helpful in practice, as columns with relatively small elements can produce stark results under some labelings.

- Plotting IRFs for the recovered shocks and attempting to name the shocks based on the dynamics is also an option, as in Brunnermeier, Palia, Sastry, & Sims (2017).

- Comparison to the historical record (events when shocks of certain size or sign should have occurred, as in Ludvigson, Ma, & Ng (2018)) can label shocks.

It is important to note that some consider the interpretation of shocks recovered using statistical identification methods (like identification via heteroskedasticity) to be a more difficult problem. Kilian & Lütkepohl (2017) argue that these shocks need not be economically meaningful. The labeling exercise outlined above does not, however, necessarily assume the shocks are meaningful - it is possible that no shock meets a theoretically-motivated labeling criterion satisfactorily. A researcher so concerned can test whether a statistically-recovered shock represents a particular economic shock by formally testing conventional identifying assumptions as overidentifying restrictions. An alternative is to informally evaluate the extent to which the impulse response functions (IRFs) align with those based on economic theory, as in Brunnermeier et al (2017) or Lütkepohl & Netšunajev (2014). While the labeling problem may be non-trivial, TVV-ID improves on existing approaches by delivering the candidate shocks under weaker assumptions.

Importantly, inference approaches that are valid for an estimated $\widehat{H}$ will also be valid for a labeled column of $\widehat{H}$, denoted $\widehat{H^{(j)}}$, under standard conditions. In general, the use of statistical measures to select a column of an estimated matrix will impact the asymptotic distribution of the ultimate column estimates. However, for all methods described that select a unique shock, the labeling criterion is consistent in the probability limit sense. This means that as $T \to \infty$, the probability of selecting the correct column based on the criterion approaches unity. Pötscher (1991) establishes asymptotic distributions in a discrete model selection setting building on intuition dating back to at least Geweke & Meese (1981). The strong notion of consistency of the labeling criterion in this context makes it direct to show that a strong form of Pötscher's results hold. This means that the if a labeling method is consistent, and the asymptotic distribution of $\widehat{H}$ is known, the selected column $\widehat{H^{(j)}}$ simply inherits that asymptotic distribution. In other words, the labeling problem can be ignored for the purpose of asymptotic inference.

A further advantage of statistical approaches to identification, which include TVV-ID, is that it is straightforward to describe the impact of economic assumptions quantitatively. Because they are used for discrete decisions – one shock or another is the policy shock – it is possible to report and compare estimated economic effects under alternative labeling assumptions. In numerous empirical applications considered, a large number of labeling assumptions agree on the policy shock. Reporting this makes a single result compelling to readers believing any of that collection of assumptions.

4 Estimators and performance

The strengths of TVV-ID is that it is an identification argument not tied to any model or estimator. It can thus be implemented by any estimator that fits an autocovariance of the residuals to the data. This can either be explicit – in the case of GMM on equation (6) – or implicit, in the case of many likelihood models. This is in contrast to the Sentana & Fiorentini or Rigobon arguments, which require either a GARCH-type model or regimes, respectively. This means that a researcher can choose a completely non-parametric approach (GMM), whatever model she thinks best describes the data (quasi-maximum likelihood, QML), or compare a variety of different models for robustness.

As noted briefly above, while identification does not require stationarity, in order for all of the estimators considered here to be well-behaved, Some degree of stationarity must be assumed. In the case of GMM, for example, ε_t must be fourth-order stationary so that the identifying moments (6) can be consistently estimated. In an SV model, $\log \sigma_t^2$ must be second-order stationary so the parameters of the SV process governing the autocovariance can be consistently estimated.

While GMM is the natural, entirely non-parametric implementation of TVV-ID, the higher moments used for identification can be very noisily estimated in realistically short macro time series. This motivates turning to likelihood approaches, which make parametric assumptions in exchange for possible efficiency gains. QML is a natural way to incorporate the identifying information of multiple autocovariances, implied by a functional form. The drawback of any likelihood-based approach is the necessity of specifying a law of motion for the structural variances; to some extent this may seem a return to parametric assumptions this paper set out to avoid. However, thanks to the general identification arguments offered above, *identification* is not tied to a particular functional form. In particular, the SV model is a common, highly flexible model of time-varying volatility very popular in the financial econometrics literature, where much work has compared its ability to describe the data with GARCH and other models (e.g. Diebold & Lopez (1995), Kim, Shephard, & Chib (1998), Barndorff-Nielsen & Shephard (2002)). There is reason to believe it could at least be a competitor to GARCH-based approaches. Bertsche & Braun (2018) adopt the model to estimate a SVAR under heteroskedasticity (without the theoretical justification offered by TVV-ID), and find it performs well in simulation. Carriero, Clark, & Marcellino (2018) use it to capture time-varying volatility in an SVAR (identification follows from particular model features), as do many Bayesian applications (e.g. Uhlig (1997), Primiceri (2005)).

In this section, I put the three heteroskedasticity-based identification schemes discussed in this paper, as well as identification based on non-Gaussianity, to the test. I consider several different implementations in simulation studies based on a wide range of DGPs. The DGPs are empirically calibrated from a bivariate SVAR (using an AR(1) SV model) where the two variables are the first principal component extracted from the McCracken & Ng FRED-MD database and the Fed Funds rate. For each further volatility process, the model is estimated for the structural shock series resulting from the AR(1) SV estimates. The H matrix used in the simulations is

$$H = \begin{bmatrix} 1 & 0.298 \\ 0.033 & 1 \end{bmatrix}.$$

Table 1: The presence of off-diagonal elements

	$E[\varepsilon_{1t}\varepsilon_{2t} \mid t \in T]$	$E[\varepsilon_{1t}\varepsilon_{2t} \mid t \in A]$	$E[\varepsilon_{1t}\varepsilon_{2t} \mid t \in B]$
$H = I_2$	-0.001	-0.002	0.000
$H = \left([1,1]', [1,1]'\right)$	-0.001	0.414	-0.415

The table computes the conditional expectations noted via simulation. The variance matrix is I_2 for 500,000 observations and $\left([1,0]', [0,2]'\right)$ for 500,000. The data is split into subsamples based on the trace of $\eta_t\eta_t'$. A is the subset of observations with trace above the median; B is A's complement.

Values of the parameters for the volatility models can be found in the Supplement.

4.1 Bias and sensitivity in the Rigobon model

First, I highlight a source of bias in the Rigobon model heretofore unaddressed in the literature. Rigobon (2003) shows that estimates are robust to misspecification of regimes, but the implicit assumption is made throughout the literature that the structural shocks are orthogonal within each regime. However, when regimes must be estimated, this is likely not the case. Since realized values of $\eta_t\eta_t'$ (and thus $\varepsilon_t\varepsilon_t'$) are used to determine regimes, besides being based on the values of σ_t, the estimated regimes are also partially driven by realized values of ε_t. Which values are conducive to a "high" vs. "low" determination will depend on the norm considered and also the value of H. This means that conditional on being in the high or low variance regime, ε_t may no longer be orthogonal, since the regimes are in effect endogenous.

A simple numerical example demonstrates this phenomenon in Table 1. 500,000 draws are taken from each of two variance regimes (I_2 and $\left([1,0]', [0,2]'\right)$) for two different H matrices. Regimes are determined from the median of $tr\left(\eta_t\eta_t'\right)$. For $H = I_2$, the shocks in each regime remain orthogonal, since the cross-term $\varepsilon_{1t}\varepsilon_{2t}$ does not impact the trace. However, for H a matrix of ones, the cross-term $\varepsilon_{1t}\varepsilon_{2t}$ matters just as much as the diagonal terms, and the shocks are strongly correlated (either positively or negatively) within regimes, while remaining orthogonal overall.

A lack of orthogonality within the sub-samples biases estimates of H. Consider the 2-dimensional case. When orthogonality holds, $E[\varepsilon_{1t}\varepsilon_{2t} \mid t \in A] = 0$, where A is one of the regimes, so

$$\sigma^2_{\eta_1,A} = E\left[\varepsilon^2_{1t} \mid t \in A\right] + H^2_{12}E\left[\varepsilon^2_{2t} \mid t \in A\right] = c_1 + H^2_{12}c_2.$$

Without orthogonality,

$$\sigma^2_{\eta_1,A} = E\left[\varepsilon^2_{1t} \mid t \in A\right] + H^2_{12}E\left[\varepsilon^2_{2t} \mid t \in A\right] + 2H_{12}E\left[\varepsilon_{1t}\varepsilon_{2t} \mid t \in A\right]$$

$$= c_1 + H^2_{12}c_2 + H_{12}c_3,$$

which includes an additional unknown, c_3. It is clear that assuming $c_3 = 0$, as the literature does, biases estimates. Alternatively, the Rigobon argument, which yields just-identification with two regimes, is now under-identified if c_3 must be determined.

This issue is clearest in the Brunnermeier et al (2017) version of the Rigobon argument. If $S_A = H\Sigma_A H'$ where $\Sigma_A \equiv E\left[\varepsilon_t\varepsilon'_t \mid t \in A\right]$ and similarly for B, then

$$S_A S_B^{-1} = H\Sigma_A\Sigma_B^{-1}H^{-1}.$$

If $\Sigma_A\Sigma_B^{-1}$ is diagonal, then those diagonal elements are the eigenvalues of the matrix on the right hand side, and the columns of H are the corresponding right eigenvectors. However, if Σ_A, Σ_B, and thus $\Sigma_A\Sigma_B^{-1}$ are not diagonal, then the diagonal elements are not the eigenvalues of the matrix, and the columns of H are not the eigenvectors.

This problem is likely to manifest if the true variance process is continuous. With many small variance changes, the determination of regimes, on the margin, can be strongly affected by off-diagonal values of Σ_t. In practice, regimes are usually estimated based on a norm of $\eta_t\eta'_t$ calculated over some rolling window. A researcher faces a trade-off: longer windows mean that the periods around regime transitions, where non-orthogonality is likely most influential, are less important. On the other hand, longer windows (particularly if there is continuous variation in volatility) can lead to weak identification (as all regimes converge to the population mean of $\eta_t\eta'_t$. A Markov switching likelihood model can explicitly require orthogonality of shocks within regimes, and is a natural solution when regimes are not known from external information. Another alternative is to explicitly accept off-diagonal terms and estimate them using additional regimes.

Finally, in light of these considerations, I illustrate the sensitivity of the Rigobon approach to the norm of $\eta_t\eta'_t$ used, the rolling window over which it is calculated, and the regime cut-off. I consider 1, 7, and 13 period rolling windows, the norms are $tr\left(\eta_t\eta'_t\right)$ and η^2_{1t} (the dimension with highest variation in volatility – approach adopted in Rigobon & Sack (2003)), and thresholds are the median and the mean plus one

Table 2: Mean estimates for Markov-switching DGP

window		1-period			7-period			13-period			oracle	
norm	threshold	H_{21}	H_{12}	RMSE	H_{21}	H_{12}	RMSE	H_{21}	H_{12}	RMSE	H_{21}	H_{12}
trace	median	0.077	-0.24	3.66	0.017	0.405	4.94	0.0021	0.382	4.45	0.033	0.288
	mean+s.d.	0.058	-0.112	6.05	0.005	0.470	6.82	0.008	0.438	6.58		
$\bar{\eta}_1^2$	median	0.061	-0.034	1.90	0.013	0.397	6.07	0.017	0.401	5.20	RMSE:	
	mean+s.d.	0.066	-0.150	4.35	0.009	0.413	7.21	0.008	0.425	6.92	2.38	

Mean estimates of estimates for Rigobon estimates for the empirically-calibrated Markov-switching DGP, $T = 200$, 5,000 draws. The window indicates the length of the rolling window over which variances were computed to form subsamples. The norm indicates the method used to evaluate the magnitude of the variance over each window. The threshold indicates the value a window had to surpass for its central observation to be considered "high variance". Estimation via Brunnermeier et al (2017) method. Labeling proceeds via an infeasible method matching H estimate to the true H to minimize L_2 norm. RMSE is weighted sum of single parameter RMSE.

standard deviation (see Rigobon & Sack (2003)). The DGP is the Markov switching model (where the regime model is well-specified). Table 2 displays the results. The oracle (based on true regimes) performs very well. Otherwise, estimates are highly sensitive to regime estimation. In general, longer windows appear to perform better. The trace seems to be a more reliable norm, with the median a better threshold. Similar results hold for an AR(1) SV process and are borne out by the full distribution of estimates (Table 11 and Figures 10-12 in the Supplement).

4.2 Performance in simulation

I now compare the relative performance of the three heteroskedasticity identification schemes plus non-Gaussianity and associated estimators across a variety of DGPs. For TVV-ID, I consider an AR(1) SV QML implementation generalizing the EM algorithm of Bertsche & Braun (2018), exploiting the expansions of Chan & Grant (2016) in the E-step. I also use a 2-step GMM estimator, making use of the first autocovariance augmented by $E[\eta_t \eta_t']$. For the Sentana & Fiorentini approach, I adopt what I refer to as the "hybrid GARCH" estimator, a GARCH(1,1) model where the autoregressive parameters are calibrated to macro data, but the mean parameters are estimated, as well as the standard GARCH(1,1) model adopted in Normandin & Phaneuf (2004) and many others. For the Rigobon approach, I use estimated regimes based on the trace, median threshold, and 13 period windows, as recommended by the results above. I also use an arbitrary split at $T/2$, as well as a Markov switching model estimated via ML. Finally, for non-Gaussianity, I use the FastICA algorithm described in Shimizu, Hoyer, Hyvärinen, & Kerminen (2006) (preliminary simulations show performance superior to ML approaches). Details on selected estimators can be found in the Supplement.

The empirically calibrated DGPs are a Markov switching model, a GARCH(1,1) model (including a "weak" variant), an AR(1) SV model (including varied sample size, a "weak" variant, non-Gaussian (t_7) disturbances), and a homoskedastic model with t_7 disturbances. I take 5000 replications, and unless otherwise noted, $T = 200$. Columns of H are labeled using the infeasible method of minimizing the norm to the true value.

Table 3 reports the results. It lists the mean estimates for the off-diagonal elements of H, RMSE (root of weighted sum of MSEs for both parameters), and rejection rates for nominal 5% tests of the true parameter values using each estimator's appropriate standard errors (described in the Supplement). Histograms reported in the Supplement show that distributions for most estimators and DGPs are centered around the true parameters; large discrepancies in mean estimates are mostly driven by outliers.

Across DGPs, the QML implementation of the AR(1) SV model performs best. The mean estimates are accurate, and even when misspecified the RMSE is often only slightly worse than that for well-specified estimators. This makes it a compelling choice to implement TVV-ID. A further benefit is that tests of true values are well-sized.

The hybrid GARCH estimator and Markov switching estimators offer the next best performance. The mean estimates are still accurate, but their RMSEs are higher in general. They struggle in the face of weak variation in volatility. For the hybrid, this this is largely because the calibrated parameters are no longer a good fit for the data. The standard errors for both estimators offer minimal size distortions.

The FastICA estimator exploiting non-Gaussianity is also reliable. The mean estimates are close to the true values except for DGPs with small sample sizes or weak variation. In these cases, the higher moments on which this identification rests seem very imprecisely estimated – moreso than the *persistence* of the process, which TVV-ID exploits. In contrast, estimators like that for SV or GARCH models exploit a path of variances for identification as well as these unconditional higher moments of the data. The RMSE is accordingly higher, depending on the DGP. Naturally, its performance improves when disturbances are themselves non-Gaussian. The standard errors perform quite poorly with respect to rejection rates – this is because the asymptotic variance depends on up to the sixth moment of the shocks, so is very imprecisely estimated.

The GARCH estimator generally is close to the previous approaches, but breaks down for SV with $T = 400$. This is because the empirical calibration dictates GARCH parameters that are very close to non-stationarity. As a result, with a longer draw of data, there is a reasonable chance of observing dynamics that appear explosive from a GARCH-fitting perspective, negatively impacting the estimates. This also appears in un-reported simulations for different empirical calibrations, generally manifest in excess mass around zero for the H parameters when the GARCH parameters are close to the boundary of stationarity. Since these calibrations are empirical, this is a strike against adopting GARCH estimators for identifying SVARs in similar macro data. The rejection rates are accurate when well-specified, but as expected, break down when misspecified.

GMM generally struggles, especially with small samples and weak variation. Since it relies mostly on higher fourth moments of identification, this makes sense, as these moments are noisily estimated in those DGPs. Accordingly, the rejection rates are also distorted. For progressively higher T, additional simulations suggest performance does become acceptable. Thus, for larger sample sizes, GMM may offer an alternative requiring no parametric assumptions.

Table 3: Mean estimates and rejection rates

		QML AR(1) SV		GMM		Hybrid		GARCH		Sub-sample (rolling)		Sub-sample ($T/2$)		Markov Switching		Non-Gaussianity	
		mean	α	mean	α	mean	α	mean	α	mean	α	mean	α	mean	α	mean	α
Markov switching, $T = 200$	H_{21}	0.03	6.8	0.01	39.9	0.02	10.0	0.03	47.2	0.02	17.0	0.01	22.6	0.03	4.1	0.03	38.5
	H_{12}	0.28	10.1	0.45	44.2	0.34	12.2	0.31	45.4	0.38	4.3	0.38	4.2	0.28	4.5	0.27	43.7
	RMSE	2.70		6.78		5.49		4.87		4.45		6.61-		2.45		4.86	
GARCH(1,1), $T = 200$	H_{21}	0.03	5.3	0.03	26.6	0.03	4.3	0.03	4.8	0.03	15.8	0.03	15.6	0.03	11.2	0.03	11.4
	H_{12}	0.29	6.8	0.40	32.0	0.33	5.3	0.30	4.7	0.19	2.3	0.36	2.8	0.37	11.1	0.41	13.3
	RMSE	2.96		7.73		2.47		2.58		5.47		6.89		5.28		6.98	
GARCH(1,1), $T = 200$, weak	H_{21}	0.03	19.0	0.02	48.7	0.02	24.4	0.03	4.8	0.03	14.1	0.02	21.6	0.03	9.6	0.02	8.1
	H_{12}	0.32	21.7	0.84	51.9	0.98	24.6	0.27	5.8	0.11	1.6	0.83	2.3	0.58	9.5	1.19	11.2
	RMSE	8.52		12.15		8.27		6.94		7.99		13.28		11.05		15.02	
AR(1), $T = 100$	H_{21}	0.03	14.9	0.01	45.2	0.03	9.2	0.03	22.3	0.02	19.4	0.02	17.0	0.02	11.5	0.02	31.2
	H_{12}	0.29	16.3	0.74	49.8	0.35	10.2	0.37	21.5	0.38	4.3	0.39	3.4	0.43	10.6	0.41	31.7
	RMSE	8.52		12.15		8.27		6.94		7.99		13.28		11.05		15.02	
AR(1), $T = 200$	H_{21}	0.03	7.8	0.02	40.6	0.03	6.9	0.03	23.4	0.02	19.1	0.02	18.9	0.03	5.4	0.03	35.4
	H_{12}	0.29	9.6	0.51	44.1	0.32	7.7	0.30	22.1	0.37	3.3	0.36	3.3	0.34	5.1	0.33	38.5
	RMSE	2.89		7.89		4.31		3.91		6.91		6.29		5.20		5.92	
AR(1), $T = 400$	H_{21}	0.03	4.2	0.02	33.7	0.03	4.3	0.06	61.7	0.02	25.2	0.02	19.9	0.03	5.5	0.03	37.1
	H_{12}	0.30	5.9	0.45	38.7	0.29	4.9	0.74	51.6	0.31	3.4	0.37	3.1	0.28	5.1	0.33	41.1
	RMSE	1.42		6.28		2.52		8.50		6.23		5.72		3.17		4.42	
AR(1), $T = 200$, weak	H_{21}	0.03	41.5	0.01	44.7	0.02	47.6	0.02	8.6	0.04	23.9	0.02	28.3	0.02	15.5	0.01	10.1
	H_{12}	0.33	42.9	0.63	45.2	0.50	48.6	0.50	10.1	0.09	2.8	0.47	3.6	0.51	14.9	0.52	12.0
	RMSE	7.63		9.08		8.36		7.80		6.38		8.41		8.00		8.96	
AR(1), $T = 200$, t_7 shocks	H_{21}	0.03	4.7	0.02	40.1	0.03	8.7	0.03	32.8	0.02	18.8	0.02	17.6	0.03	4.0	0.03	32.4
	H_{12}	0.30	5.8	0.60	44.0	0.32	9.6	0.31	30.7	0.33	3.6	0.38	3.6	0.32	3.5	0.30	34.4
	RMSE	2.24		8.45		5.62		4.95		6.38		6.27		5.65		4.54	
homosked. $t_7$shocks	H_{21}	0.03	5.8	0.01	39.3	0.00	66.4	0.01	4.6	0.05	20.2	0.01	26.3	0.02	19.8	0.03	16.7
	H_{12}	0.28	5.4	0.43	40.0	0.50	67.5	0.39	4.9	0.08	2.9	0.40	3.9	0.34	18.1	0.30	17.3
	RMSE	3.07		6.78		8.16		7.06		4.23		6.93		5.45		3.88	

Mean estimates for the full range of estimators for the specified DGPs. Labeling proceeds via an infeasible method matching H estimates to the true H to minimize L_2 norm. Rejection rates, α, are presented for a nominally-sized 5% test for each draw. Details on standard errors can be found in the Supplement. Since the RMSE must account for error in multiple parameter estimates, the MSE is computed for each, and then normalized by the square of the true parameter, before the root of the sum is taken.

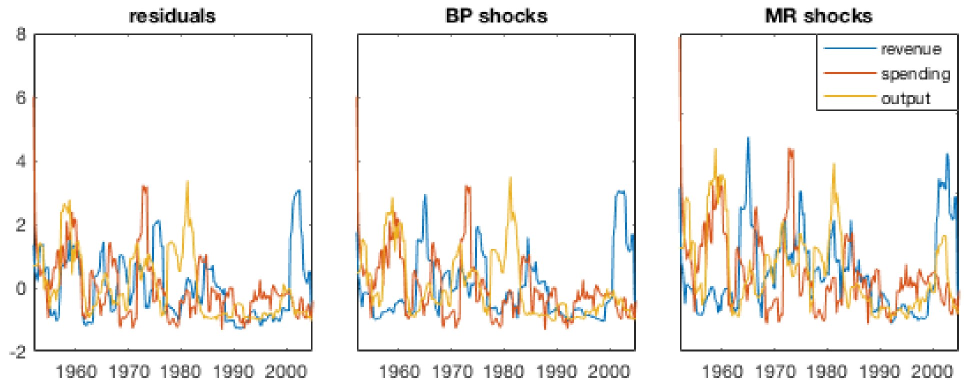
The Rigobon estimates based on rolling windows are quite good, which is unsurprising given the tuning parameters were optimized based on Section 4.1. Other combinations might dramatically harm performance. However, the breakdown is dramatic for weak identification, which makes sense as that is when the argument for bias presented in 4.1 is most impactful. However, it is not in general competitive with the best estimators in terms of RMSE. Generally, the same remarks apply to the simple $T/2$ split estimator. For both, the rejection rates are badly distorted.

The homoskedastic DGP with non-Gaussian shocks warrants a separate note. The SV model estimates remain reliable (indeed better than FastICA in terms of RMSE). Seeing as all estimators but FastICA rely on heteroskedasticity for identification, this may be surprising. Recall that the SV estimator exploits a variance path for identification. Even though there is no variation in the true structural variances, the t_7 distribution of shocks with constant volatility can be approximated by a Gaussian model with time-varying volatility taking higher values when tail-draws are observed. The GARCH estimator, while in principle similar, cannot provide as good an approximation, as innovations are restricted to be a multiple of previous shocks. The finding that the SV estimates can be reliable even under homoskedasticity provides a further commendation for its use in applied work. It effectively simultaneously harnesses both heteroskedasticity and non-Gaussianity for identification in a way other methods do not.

5 Empirical application: fiscal multipliers

Considerable work has been devoted to estimating the value of fiscal multipliers, but has resulted in considerable disagreement over their size. The range of estimates is documented by Mertens & Ravn (2014), Caldara & Kamps (2017), and Ramey (2011a). Prominent estimates range from less than zero to over three. While government spending multipliers are perhaps most familiar, tax multipliers capture an equally important dimension of fiscal policy, and are central to current policy debates. Blanchard & Perotti (2002) (henceforth BP) is seminal in the literature; recent work by Mertens & Ravn (2013, 2014) and Mountford & Uhlig (2009) has obtained contrasting estimates. Caldara & Kamps (2017) show the discrepancy can be largely explained by differing values for the elasticity of tax revenues with respect to output. BP calibrate this parameter to 2.08 based on institutional information, Mountford &

Figure 2: Moving averages of squared residuals and shocks



2-year moving averages of the square of the specified series. For the first panel, this is the reduced form residuals, BP structural shocks for the second, and MR structural shocks for the third.

Uhlig's (2009) penalty-function identification is consistent with a prior for the elasticity centered around 3, and Mertens & Ravn (2014) (henceforth MR) estimate a value of 3.13 using Romer & Romer (2010) (henceforth RR) narrative shocks as external instruments.

Theorem 1 shows that the autocovariance of volatility present in the data can identify the structural parameters determining fiscal multipliers without the economic assumptions required in prior work. BP need a calibrated value, there to be no contemporaneous response of spending to output, and a recursive ordering between tax revenue and spending, and MR require their instrument to be valid and there to be no contemporaneous response of spending to output. Since I am able to depart from these assumptions (making them over-identifying restrictions), I can test them using the results of TVV-ID. To motivate TVV-ID, Figure 2 plots "eyeball" evidence of heteroskedasticity in the data using moving averages of squared disturbances for the reduced form residuals, BP's shocks, and MR's shocks in turn. Formal non-parametric tests for the *autocovariance* of the structural shocks (as required by Theorem 1) are difficult, requiring identification to hold without time-varying volatility. Applying a suitable test proposed by Lanne & Saikkonen (2007) to BP's and MR's shocks rejects the null of no autocovariance in the volatilities at the 5% level for a majority of bandwidths. This is reasonable evidence, since Lütkepohl & Milunovich (2016) find the test to have very low power; details can be found in the Supplement.

5.1 Data & model

I adopt MR's trivariate VAR with federal tax revenue, federal government consumption and investment, and GDP, based on quarterly BLS data found in the NIPA tables, spanning 1950Q1 to 2006Q4. Additional details on the data and de-trending procedures (including federal vs. general government data) can be found in MR. I use the replication code available on Mertens' website to obtain identical residuals.

In MR's notation, the BP benchmark model is

$$\begin{aligned} u_t^T =& \sigma_T e_t^T + \theta_G \sigma_G e_t^G + \theta_Y u_t^Y \\ u_t^G =& \gamma_T \sigma_T e_t^T + \sigma_G e_t^G + \gamma_Y u_t^Y \\ u_t^Y =& \zeta_T u_t^T + \zeta_G u_t^G + \sigma_Y e_t^Y , \end{aligned}$$

where $u_t = \eta_t$ and e_t are structural shocks with $E\left[e_t^{'} e_t\right] = I$. Key parameters θ_Y and γ_Y are the elasticities of tax revenue and government spending with respect to output, respectively. These capture what are commonly referred to as "automatic stabilizer" effects. This model is a transformation of the $\eta_t = H\varepsilon_t$ parameterization. The transformations linking the parameters to H are

$$\begin{aligned} \theta_G &= \frac{H_{12} - H_{32}H_{13}}{1 - H_{23}H_{32}}, \theta_Y = H_{13} \\ \gamma_T &= \frac{H_{21} - H_{23}H_{31}}{1 - H_{31}H_{13}}, \gamma_Y = H_{23} \\ \zeta_T &= \frac{H_{31} - H_{32}H_{21}}{1 - H_{21}H_{12}}, \zeta_G = \frac{H_{32} - H_{31}H_{12}}{1 - H_{21}H_{12}}, \end{aligned} \tag{8}$$

This mapping allows for direct comparison with the TVV-ID results.

5.2 Estimates & tests

Estimation based on TVV-ID proceeds using the AR(1) SV approach recommended by the simulation study. The estimates are reported in the third row of Table 4, with BP and MR results for comparison.[8] The structural shocks themselves are extremely well-correlated with the BP shocks and very well-correlated with the MR

[8]It is well-known that EM algorithms can be sensitive to start values; thus, optimization was carried out across a grid of start values, and the median estimates used to initialize a final optimization. The range of estimates across start values is very small, see Table 5 in the Supplement. As an additional check, the estimates from a Markov switching model (same Table) are extremely similar.

Table 4: Estimates

	BP	MR	TVV-ID
θ_G	-0.06	-0.20	-0.13 (0.10)
θ_Y	2.08	3.13	1.58 (0.18)
γ_T	0	0.06	0.11 (0.13)
γ_Y	0	0	0.02 (0.39)
ζ_T	-0.08	-0.35	-0.00 (0.02)
ζ_G	0.07	0.10	0.06 (0.045)

The first two columns are estimates obtained in Mertens & Ravn (2014). The third column maps estimates of H obtained via TVV-ID to the parameters of BP and MR using (8). The TVV-ID estimates result from fitting the AR(1) SV model, described in Section 4, with details provided in Supplement 2.2.

shocks. The one statistically significant parameter estimate is that central to the tax multiplier debate, θ_Y, for which I obtain the value 1.58 with a 95% confidence interval of $[1.23, 1.94]$. The estimated SV parameters offer parametric evidence of the autocovariance exploited by TVV-ID in Theorem 1. The AR parameters are 0.71 ($p = 0.002$), 0.88 ($p = 0.085$), and 0.93 ($p = 6 \times 10^{-10}$) for tax, spending, and output shocks respectively. The innovations to the log-variance are imprecisely estimated, but are on the same order as those in the simulation study.[9] Figure 6 in the Supplement plots the filtered means of the variance paths, which exhibit considerable heteroskedasticity.

Testing the Blanchard & Perotti (2002) assumptions

The three identifying assumptions made by BP can be directly tested from the estimates of H. First, for the elasticity of tax revenues with respect to output, θ_Y, I obtain a value of 1.58, and can reject BP's calibrated value 2.08 at the 1% level. In the version of their model documented in MR, spending is assumed to respond

[9] Identification of the shocks (and thus the precise decomposition of overall autocovariance into the three series) is conditional on time-varying volatility, but the overall autocovariance properties these estimates decompose are independent of the identification of H.

contemporaneously only to its own shocks: $\gamma_T = \gamma_Y = 0$. In the original paper, $\theta_G = 0$ (taxes do not respond to spending) is an alternative to $\gamma_T = 0$. None of these exclusion restrictions can be rejected; they are consistent with TVV-ID results.

Testing the validity of Mertens & Ravn's (2014) instruments

MR use the RR shocks as external instruments to identify tax shocks. Like standard instruments, they must be both relevant and exogenous (see Montiel Olea, Stock, & Watson (2016)). Thus, for relevance, I compute first-stage $F-$statistics under both homoskedasticity and heteroskedasticity, and compare them to the $F > 10$ rule of thumb of Staiger & Stock (1997) and the corresponding $F > 23$ rule following Montiel Olea & Pflueger (2013). Under homoskedasticity the value is 4.13 and under heteroskedasticity 1.76; the instrument is only weakly related to the endogenous residual. This suggests there could be a weak identification problem. Table 6 in the Supplement shows that this is true of all alternative narrative measures considered by MR. These results are at odds with the reliability measure they report. This measure of how much variation in the instrument is explained by the structural shock is asymptotically equivalent to the R^2. There are reasons to favour conclusions based on the first-stage $F-$statistic. The reliability measure can only be computed based on estimated structural shocks; instrument validity is *assumed* to obtain these. The $F-$statistic also conveys more information because established thresholds are based on how a deficiency in the first-stage quantitatively impacts bias or size-distortion in the second stage.[10]

Using the structural shocks from TVV-ID, I can also test the exogeneity assumption required for the proxy VAR. I test the hypothesis that the coefficients in the regression of the RR shocks on ε_t^G and ε_t^Y are zero. The test rejects at the 5% level for the shocks jointly, driven by a significant negative relationship with ε_t^Y. This suggests that, despite careful construction, the narrative measure has not been fully purged of cyclical behaviour, and still contains endogenous variation in tax revenues. Table 6 in the Supplement repeats the exercise for the alternative shocks in MR; only

[10]Additionally, MR note that the reliability statistic requires the additive form of measurement error specified in the text. However, it is reasonable to believe measurement error could scale with the size of the tax shock being measured, in keeping with several common forms of heteroskedasticity (in linear regression). The reliability itself also offers no measure of the uncertainty around the relationship between the shocks and instrument. While MR do bootstrap the statistic, it is well-known that bootstrapping procedures may not properly capture variability if weak identification is present.

series based on the full set of RR shocks (including shocks with implementation lags) do not exhibit endogeneity. The strong negative relationship between the instrument and output shocks implies that, for a tax cut, the estimated impact on output could be biased upwards.

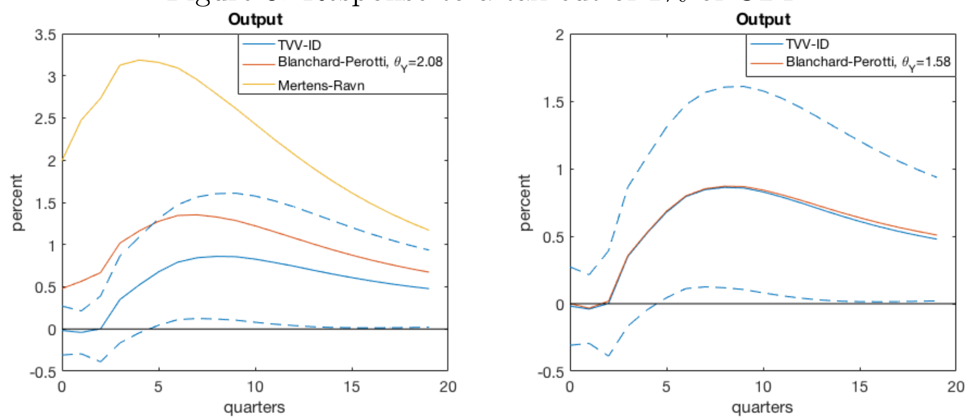
5.3 Multipliers

The parameter estimates of TVV-ID lead to important differences in dynamic multipliers compared to previous work. Figure 3 plots the dynamic tax multiplier following the methodology of MR. The shock corresponds to a tax cut of 1% of GDP. 95% confidence intervals are computed using the same wild bootstrap as MR for the reduced form portion of the IRF with the ML variance estimates of the structural parameters, combined using the delta method. The differences compared to the BP and particularly MR results in the first panel are stark. The MR IRF is rejected at all horizons; the BP for horizons up to five quarters. As discussed above, endogeneity of the RR shocks with respect to output shocks could be causing an upward bias for MR. The response on impact is -0.02% (not significant) compared to 0.48% for BP and 1.99% for MR. The peak multiplier occurs later and is lower: 0.86 at eight quarters compared to 1.35% at seven quarters (BP) and 3.19% at four quarters (MR). It suggests a more significant response lag of the economy to tax changes than previous results. The second panel recomputes the IRF for BP using the new elasticity estimated via TVV-ID. The path is virtually identical to the TVV-ID path, mimicking the result when MR do the same using their estimated elasticity. This affirms the finding of Caldara & Kamps (2017) that the elasticity explains virtually all estimated differences in multipliers, and shows that the results of BP can be reconciled with those of TVV-ID via the calibrated parameter.

Figure 4 plots the government spending multipliers. The estimates here are much more similar across approaches, as predicted by more similar values for θ_G and ζ_G. On impact, the multipliers are 0.65% for TVV-ID, just lower than BP (0.69%) and close to MR (0.80%). The maximum response is 0.75% for TVV-ID, compared to 0.81% for BP and 0.96% for MR, all after two quarters. The second panel plots the BP response with the elasticity re-calibrated to TVV-ID; doing so barely impacts estimates of $H^{(G)}$, so the paths are virtually identical.

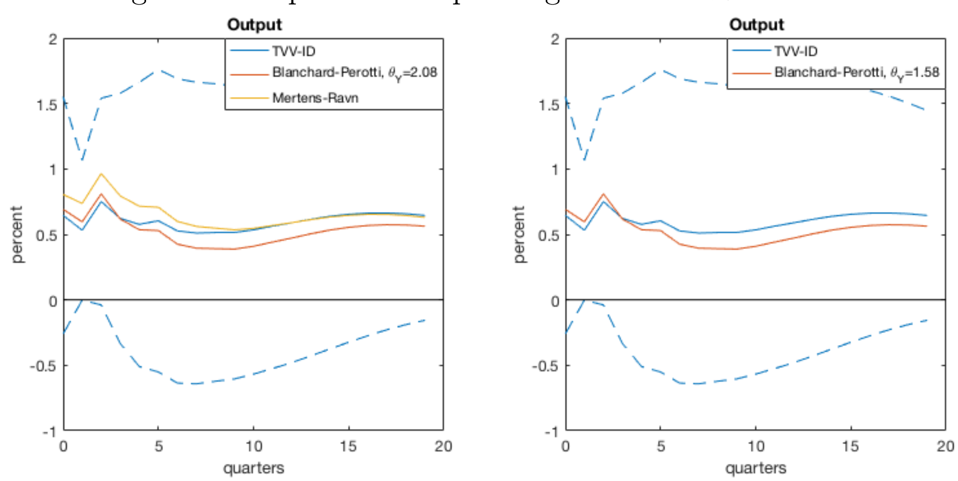
Caldara & Kamps (2017) develop a new methodology using non-fiscal proxies as

Figure 3: Response to a tax cut of 1% of GDP



Dashed lines are 95% confidence intervals computed using a wild bootstrap for the reduced form and ML for the structural parameters, combined using the delta method. The BP estimates in the left panel use their elasticity $\theta_Y = 2.08$; the right uses the value of 1.58 estimated via TVV-ID.

Figure 4: Response to a spending shock of 1% of GDP



Dashed lines are 95% confidence intervals computed using a wild bootstrap for the reduced form and ML for the structural parameters, combined using the delta method. The BP estimates in the left panel use their elasticity $\theta_Y = 2.08$; the right uses the value of 1.58 estimated via TVV-ID.

instruments for identification and find that, in the short run, spending multipliers are larger than tax multipliers. This is also true here, but tax multipliers do eclipse spending multipliers around their peak impact, after two years. Figure 7 in the Supplement plots their estimates agains mine. In general, the IRFs are similar, but TVV-ID yields smaller multipliers on impact. Except for quarters 0-4 for tax cuts, my point estimates lie within their 68% credible sets; all estimates lie within 95% credible sets. They also find that the impact effect largely explains discrepancies in dynamic multipliers across identification approaches, and in turn, that the impacts are governed almost entirely by the elasticities with respect to output and the covariances of u_t (their equation (11)). Accordingly, for their version of BP, using a tax elasticity of 1.7, they obtain an almost identical IRF to mine. For spending, the slightly higher (though not statistically significantly so) spending elasticity I obtain can explain the somewhat lower multipliers obtained. Differences in the shape of the IRFs result from different reduced form coefficients (they use a 5-variable VAR and only a linear time trend). Since the identifying conditions of both schemes hold up to testing (their instruments pass validity tests), it is reassuring that, with the exception of taxes on impact, the dynamic multipliers of each study cannot be rejected under the methodology of the other.

The results under TVV-ID are also in line with the spending multipliers obtained by Ramey (2011b) using a very different methodology based on isolating defense-related spending events. For a sample from 1939-2008, excluding WWII, her estimates range from 0.6-0.8, which includes the values estimated here. Ramey & Zubairy (2018) consider how the multiplier changes across states of the economy using defense spending and BP spending shocks as instruments, which they check carefully for relevance. While in some relevant states of the economy they obtain somewhat lower estimates, the 0.6-0.7 range found here accords with their results.

MR are skeptical of the results of RR and Favero & Giavazzi (2012), both of which take a narrative approach, regressing on the RR tax shocks directly. They argue that in general they are likely to underestimate the true multiplier. Figure 8 in the Supplement compares my results to those of RR and Favero & Giavazzi (2012). While the RR dynamic multipliers are dramatically different, the Favero & Giavazzi results are contained in my 95% confidence set after the first year. My results on exogeneity of the RR shocks suggest that these papers may *overestimate* multipliers. If the weak relevance results are caused by the series only picking up the most prominent episodes, that would offer another channel for the estimates to be biased upwards.

The elasticity of revenues with respect to output

The differences in tax multipliers are largely determined by the lower elasticity of tax revenues with respect to output that I estimate via TVV-ID. This discrepancy between my elasticity and the original BP paper may be partially explained by the fact that, in their calibration, they consider data on general government revenue and spending, as opposed to federal government. The response of federal revenues should be lower than that of federal, state, and local revenues *combined*. Significantly, Follette & Lutz (2010) develop a more detailed methodology and estimate the elasticity of tax revenues with respect to output for just the federal government, and obtain a value of 1.6 for the period 1986-2008 – nearly identical to what I obtain via TVV-ID. They obtain 1.4 for 1960-1985. On average, their value is thus slightly lower than mine, but they consider a mix of annual data and quarterly data. This accords with BP's argument that lower frequency data will deliver lower elasticities. MR discuss discrepancies between institutional estimates of BP, Giorno, Richardson, Roseveare, & van der Noord (1995), and others and results coming from their instrumental approach; such discrepancies seem less pronounced compared to TVV-ID.

The higher elasticity of MR, 3.13, may result from the weakness of the instrument, as discussed above. In more recent work, Caldara & Kamps' (2017) elasticity estimated baseline using all of their non-fiscal instruments is lower at 2.18. Their instruments pass all tests for relevance and exogeneity, unlike the RR shocks. Their elasticity is still higher than mine, but is based on a different reduced-form VAR. They also show that the high elasticity found in Mountford & Uhlig (2009) – about 3.2 – can be traced to their penalty-function identification approach, which *maximizes* the systematic component of tax revenues.

6 Conclusion

This paper presents a general argument that structural shocks can be identified via time-varying volatility. The previous literature offers identification arguments based on a path of variances available for only very few parametric models of the variance process. My identification approach makes minimal assumptions on the variances as a stochastic process. This argument highlights a novel channel of identification based on heteroskedasticity that frees the researcher from needing to assume a particular functional form (or, indeed, any functional form) to obtain identifying moments. This empowers researchers to develop new models and approaches in contexts exhibiting time-varying volatility without needing to re-establish identification for each. Economic information usually used to identify such models need only be used to label the shocks identified by TVVID. A variety of estimation methods are proposed. Simulation evidence shows that quasi-likelihood methods based on an auto-regressive log-variance process work well even when the true process has a different form.

My empirical investigation of fiscal multipliers produces estimates that are quite low, but broadly align with previous studies. The tax multipliers estimated by Blanchard & Perotti (2002) can easily be reconciled with TVV-ID by adjusting their calibrated elasticity of revenues. The tax elasticity I obtain, about 1.6, is consistent with the work of Follette & Lutz (2010). For both tax changes and government spending, my results are fairly similar to those recently obtained by Caldara & Kamps (2017). For government spending, my multipliers are consistent with the values in Ramey (2011b) and Ramey & Zubairy (2018). Mertens & Ravn's (2014) high values might result from instrument endogeneity or weakness. These findings contribute to an increasing body of empirical work in favour of multipliers below unity, and to tax multipliers smaller than spending multipliers. This demonstrates the potential of TVV-ID to offer new insights into old problems using an identification approach radically different from those previous considered in the literature.

A Notation

The following potentially unfamiliar notation is used in the paper. $\otimes$ represents the Kronecker product of two matrices; $\odot$ represents the element-wise product of two matrices (i.e. Hadamard product); $A_{(i)}$ denotes the i^{th} row of matrix A; $A^{(j)}$ denotes the j^{th} column of matrix A; A_{ij} denotes the ij^{th} element of matrix A; $A^{(-i)}$ denotes all columns of A except for the i^{th}, and similarly for rows and elements; $matdiag\,(A)$ is a vector of the diagonal elements of the square matrix A; $diag\,(a)$ is a diagonal matrix with the vector a on the diagonal; $x_{1:t}$ denotes $\{x_1, x_2, \ldots, x_t\}$; $E_t\,[\cdot]$ denotes a time-specific expectation, i.e. the mean value of x_t at time t, as opposed to across t, and similarly for $E_{t,s}\,[\cdot]$ when both time t, s variables are contained in the argument.[11]

B Proofs

B.1 Derivation of Proposition 1

Proof. I start with

$$E_{t,s}\left[\zeta_t \mid \sigma_t, \mathcal{F}_{t-1}\right] = L\left(H \otimes H\right) G\sigma_t^2.$$

Since v_t was shown to be a martingale difference sequence and $\mathrm{Var}_t\left(v_t\right) < \infty$ (Assumption C.2),

$$\mathrm{Cov}_{t,s}\left(v_t, v_s\right) = 0, \;\; s \neq t.$$

This also implies that in the signal-noise decomposition, (5), v_t is white noise. Using this fact, Assumption B, Assumption C.1-2, and the decomposition of ζ_t above, it is immediate that, for $s \neq t$,

$$\begin{aligned} E_{t,s}\left(\zeta_t\zeta_s'\right) &= L\left(H \otimes H\right) G E_{t,s}\left[\sigma_t^2\sigma_s^{2'}\right] G'\left(H \otimes H\right)' L' \\ &\quad + L\left(H \otimes H\right) G E_{t,s}\left[\sigma_t^2 v_s'\right] + E_{t,s}\left[v_t\sigma_s^{2'}\right] G'\left(H \otimes H\right)' L'. \end{aligned} \tag{9}$$

[11] This notation is used to make explicit that stationarity is not being assumed, unless otherwise noted, and to avoid the ambiguity (and possible non-existence) present in simply writing $E\,[x_t]$ in a non-stationary context. The use of E_t should not be confused with reference to the t information set; when a specific information set is intended, I condition on it explicitly.

By the law of iterated expectations, Assumption A.1 implies that

$$E_{t,s}\left[\Sigma_t \mid \sigma_s^2\right] = E_{t,s}\left[\varepsilon_t\varepsilon_t' \mid \sigma_s^2\right],\ t \geq s.$$

This, in turn, by the law of iterated expectations, implies that

$$E_{t,s}\left[vec\left(\varepsilon_t\varepsilon_t' - \Sigma_t\right)\sigma_s^{2'}\right] = 0,\ t \geq s.$$

Thus, setting $t > s$, the third term in (9) vanishes, leaving

$$E_{t,s}\left(\zeta_t\zeta_s'\right) = L\left(H\otimes H\right)GE_{t,s}\left[\sigma_t^2\sigma_s^{2'}\right]G'\left(H\otimes H\right)'L' + L\left(H\otimes H\right)GE_{t,s}\left[\sigma_t^2 v_s'\right]. \quad (10)$$

Finally, I can rewrite (10) as

$$L\left(H\otimes H\right)\left(GE_{t,s}\left[\sigma_t^2\sigma_s^{2'}\right]G' + GE_{t,s}\left[\sigma_t^2 vec\left(\varepsilon_s\varepsilon_s' - \Sigma_s\right)\right]\right)\left(H\otimes H\right)'L'$$

$$= L\left(H\otimes H\right)GM_{t,s}\left(H\otimes H\right)'L' \quad (11)$$

where $M_{t,s} = E_{t,s}\left[\sigma_t^2\sigma_s^{2'}\right]G' + E_{t,s}\left[\sigma_t^2 vec\left(\varepsilon_s\varepsilon_s' - \Sigma_s\right)'\right]$. $M_{t,s}$ is an $n \times n^2$ matrix. Proposition 1 then follows directly. □

B.2 Proof of Theorem 1

I begin by proving two lemmas for properties of the singular value decomposition (SVD).

Definition 1. Define

- $U_1 D_U U_2' = V$, a reduced SVD, V $n_1 \times n_2$, D_U $d \times d$,
- C_i is a full rank matrix, $m_i \times n_i, m_i \geq n_i$,
- $F = C_1 V C_2'$, non-defective.

Lemma 1. *There exists a matrix Γ_1 such that $CU_1\Gamma_1$ is an orthogonal matrix of singular vectors from a SVD of F.*

Proof. Define $Q_1 R_1 = CU_1$, a reduced QR decomposition, and similarly for CU_2. Then $F = Q_1 R_1 D_U R_2' Q_2'$. R_1 is $d \times d$ and full rank since, given CU_1 is full rank d, it has no zeros on the diagonal (Trefethen & Bau (1997), Exercise 7.5). Now define $W_1 D_R W_2' = R_1 D_U R_2'$, another SVD; then $F = (Q_1 W_1) D_R (W_2' Q_2')$ is a reduced SVD (it is easily shown D_R are singular values of F, and the corresponding vectors are clearly orthogonal). Thus write $Q_1 R_1 \left(R_1^{-1} W_1\right) = Q_1 W_1$ so $\Gamma_1 = R_1^{-1} W_1$, which is guaranteed to exist. □

Definition 2. Define

- $S_1 D_S S_2' = F$, a reduced SVD

Lemma 2. *The SVD of F is unique up to rotations characterized by $F = S_1 T_1 D_S T_2 S_2'$ where T_i is orthogonal.*

Proof. The singular values, D_S, are unique, singular vectors corresponding to non-repeated values are unique up to sign, and the space of vectors corresponding to k repeated singular values corresponds to linear combinations of any k such vectors. Thus $F = (S_1 T_1) D_S (T_2 S_2')$ characterizes any reduced SVD as T_i can incorporate any such sign changes or linear combinations. Since $S_i T_i$ must be orthogonal, $T_i' S_i' S_i T_i = I_d$. Then since S_i is orthogonal, $T_i' T_i = I_d$, so T_i is orthogonal. □

Definition 3. Define

- In particular, $C_1 = (H \otimes H) G$, $n^2 \times n$ with rank n,
- G is a selection matrix such that $vec(ADA') = (A \otimes A) G diag(D)$,
- $\hat{S}_1 = C_1 U_1 \Gamma_1 T_1$, an arbitrary reduced SVD of F,
- V is $n \times n^2$ and has no scalar multiple rows,
- $rank(V) \geq 2$.

Proposition 4. *H is uniquely determined from the equations $F = C_1 V C_2'$ provided V has no scalar multiple rows.*

Proof. U_1 is $n \times d$. Note $CU_1 = \left[vec\left(Hdiag\left(U_1^{(1)}\right)H'\right), \ldots, vec\left(Hdiag\left(U_1^{(d)}\right)H'\right)\right]$, where $d \geq 2$. By the scalar multiples condition on V, for any column j of H, there exists at least one pair k, l such that $U_{1,j}^{(l)}/U_{1,i}^{(l)} \neq U_{1,j}^{(k)}/U_{1,i}^{(k)}$ for all $i = 1, 2, ..., d,\ i \neq j$. By an argument due to Brunnermeier et al (2017), $H^{(j)}$ is unique up to scale and sign as the right eigenvector of $Hdiag\left(U_1^{(l)}\right)H'\left(Hdiag\left(U_1^{(k)}\right)H'\right)^{-1}$ corresponding to the j^{th} eigenvalue. The same argument applies to $C\tilde{U}_1$ where $\tilde{U}_1 = U_1\Gamma_1 T_1$, provided $\tilde{U}_1$ has no scalar multiple rows. To establish this, take any two rows in U_1 that are not scalar multiples; multiplying by full-rank Γ_1 cannot decrease their rank (so they do not become scalar multiples). The same holds true for multiplication by the orthogonal T_1. Thus H remains the unique solution to $C\tilde{U}_1$. □

Proposition 4 is re-written in economic terms to yield Theorem 1.

B.3 Proof of Corollary 1

Proof. Corollary 1 follows directly from Proposition 4 above for any column j for which a pair k, l exists such that $U_{1,j}^{(l)}/U_{1,i}^{(l)} \neq U_{1,j}^{(k)}/U_{1,i}^{(k)}$ for all $i = 1, 2, ..., d$. □

B.4 Proof of Theorem 2

Proof. Theorem 2 is based on the argument underlying Proposition 4. Note that the vectorization of $E_t[\zeta_t]$ is given by $vech\,(HE_t[\Sigma_t]\,H')$, an additional equation of the form found in CU_1. Define $U_{1,M}D_M U_{2,M}' = M_{t,s}$ and $\hat{M} = \left[\begin{array}{cc} U_{1,M} & E_t[\sigma_t^2] \end{array}\right]$. Then there is an additional column over which to find a k, l pair for j such that $\hat{M}_j^{(l)}/\hat{M}_i^{(l)} \neq \hat{M}_j^{(k)}/\hat{M}_i^{(k)}$ for all $i = 1, 2, ..., d$ $i \neq j$. The condition on $M_{t,s}$ (V in Proposition 4) guaranteeing this is augmented to require no scalar multiple rows in $\left[\begin{array}{cc} M_{t,s} & E_t[\sigma_t^2] \end{array}\right]$. Note that this logic can be extended to adding additional autocovariances, etc., in each case making the length of the rows that must not be scalar multiples longer and thus decreasing the plausibility of the condition failing. □

References

Andrews, D. W. K. (1983): "First Order Autoregresssive Processes and Strong Mixing," manuscript.

——— (1993): "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica*, 61, 821–856.

Angelini, G., E. Bacchiocchi, G. Caggiano, and L. Fanelli (2018): "Uncertainty Across Volatility Regimes," Manuscript.

Barndorff-Nielsen, O. E. and N. Shephard (2002): "Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society*, 253–280.

Barro, R. J. and G. Liao (2017): "Options-Pricing Formula with Disaster Risk," manuscript.

Bernanke, B. S., J. Boivin, and P. Eliasz (2005): "Measuring the Effects of Monetary Policy: a Factor-Augmented Vector Autoregressive (FAVAR) Approach," *The Quarterly Journal of Economics*, 120, 387–422.

Bertsche, D. and R. Braun (2018): "Identification of Structural Vector Autogregressions by Stochastic Volatility," .

Bharadwaj, P., L. Demanet, and A. Fournier (2017): "Deblending Random Seismic Sources via Independent Component Analysis," manuscript.

Blanchard, O. and R. Perotti (2002): "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics*, 117, 1329–1368.

Blanchard, O. and D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79, 655–673.

Blanchard, O. J. and J. A. Simon (2001): "The Long and Large Decline in U.S. Output Volatility," *Brookings Papers on Economic Activity*, 135–174.

Blanco, D. and B. Mulgrew (2005): "ICA in Signals with Multiplicative Noise," *IEEE Transactions on Signal Processing*, 53, 2648–2657.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77, 623–685.

Brunnermeier, M., D. Palia, K. A. Sastry, and C. A. Sims (2017): "Feedbacks: Financial Markets and Economic Activity," manuscript, 1–50.

Caldara, D. and C. Kamps (2017): "The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers," *Review of Economic Studies*, 84, 1015–1040.

Campbell, J. Y., S. Giglio, C. Polk, and R. Turley (2017): "An Intertemporal CAPM with Stochastic Volatility," *Journal of Financial Economics (forthcoming)*.

Cappé, O., E. Moulines, and T. Rydén (2005): *Inference in Hidden Markov Models*, New York: Springer.

Carriero, A., T. E. Clark, and M. Marcellino (2017): "Endogenous Uncertainty?" manuscript.

Chan, J. C. C. and A. L. Grant (2016): "On the Observed-Data Deviance Information Criterion for Volatility Modeling," *Journal of Financial Econometrics*, 14, 772–802.

Davidson, J. (1994): *Stochastic Limit Theory: An Introduction for Econometricians*, Oxford: Oxford University Press.

Diebold, F. X. and J. A. Lopez (1995): "Modeling Volatillity Dynamics," in *Macroeconometrics: Developments, Tensions and Prospects*, ed. by K. D. Hoover, New York: Springer, 427–472.

Favero, C. and F. Giavazzi (2012): "Measuring Tax Multipliers: The Narrative Method in Fiscal VARs," *American Economic Journal: Economic Policy*, 4, 69–94.

Follette, G. and B. Lutz (2010): "Fiscal Policy in the United States: Automatic Stabilizers, Discretionary Fiscal Policy Actions, and the Economy," .

Fuh, C.-D. (2006): "Efficient Likelihood Estimation in State Space Models," *Annals of Statistics*, 34, 2026–2068.

Genon-Catalot, V., T. Jeantheau, and C. Larédo (2000): "Stochastic Volatility Models as Hidden Markov Models and Statistical Applications," *Bernoulli*, 6, 1051–1079.

Geweke, J. and R. Meese (1981): "Estimating Regression Models of Finite but Unknown Order," *International Economic Review*, 22, 55–70.

Giorno, C., P. Richardson, D. Roseveare, and P. van den Noord (1995): "Estimating Potential Output, Output Gaps and Structural Budget Balances," .

Gouriéroux, C. and A. Monfort (2015): "Revisiting Identification and Estimation in Structural VARMA Models," manuscript.

Gouriéroux, C., A. Monfort, and J.-P. Renne (2017): "Statistical inference for independent component analysis : Application to structural VAR models," *Journal of Econometrics*, 196, 111–126.

Hyvärinen, A., J. Karhunen, and E. Oja (2001): *Idedpendent Component Analysis*, New York: John Wiley & Sons.

Hyvärinen, A., K. Zhang, S. Shimizu, and H. P. O. (2010): "Estimation of a Structural Vector Autoregression Model Using Non-Gaussianity," *Journal of Machine Learning Research*, 11, 1709–1731.

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring Uncertainty," *American Economic Review*, 105, 1177–1216.

Kilian, L. (2009): "Not All Oil Price Shocks are Alike: Disentangling Supply Shocks in the Crude Oil Market," *The American Economic Review*, 99, 1053–1069.

Kilian, L. and H. Lütkepohl (2017): *Structural Vector Autoregressive Analysis*, Cambridge: Cambridge University Press.

Kilian, L. and C. Park (2009): "The Impact of Oil Price Shocks on the U.S. Stock Market," *International Economic Review*, 50, 1267–1287.

Kim, S., N. Shephard, and S. Chib (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *The Review of Economic Studies*, 65, 361–393.

Lanne, M. and H. Lütkepohl (2010): "Structural Vector Autoregressions with Nonnormal Residuals," *Journal of Business and Economic Statistics*, 28, 159–168.

Lanne, M., H. Lütkepohl, and K. Maciejowska (2010): "Structural Vector Autoregressions with Markov Switching," *Journal of Economic Dynamics and Control*, 34, 121–131.

Lanne, M. and P. Saikkonen (2007): "A Multivariate Generalized Orthogonal Factor GARCH Model," *Journal of Business & Economic Statistics*, 25, 61–75.

Leeper, B. E. M., A. W. Richter, and T. B. Walker (2012): "Quantitative Effects of Fiscal Foresight," *American Economic Journal: Economic Policy*, 4, 115–144.

Lindner, A. M. (2009): "Stationarity, Mixing, Distributional Properties and Moments of GARCH(p,q)-Processes," in *Handbook of Financial Time Series*, ed. by T. Mikosch, J.-P. Kreiß, R. A. Davis, and T. G. Andersen, Berlin: Springer, 43–69.

Ludvigson, S. C., S. Ng, and S. Ma (2018): "Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response?" manuscript.

Lütkepohl, H. and G. Milunovich (2016): "Testing for Identification in SVAR-GARCH Models," *Journal of Economic Dynamics and Control*, 73, 241–258.

Lütkepohl, H. and A. Netšunajev (2014): "Disentangling Demand and Supply Shocks in the Crude Oil Market: How to Check Sign Restrictions in Structural VARs," *Journal of Applied Econometrics*, 29, 479–496.

——— (2017): "Structural Vector Autoregressions with Smooth Transition in Variances: The interaction between US monetary policy and the stock market," *Journal of Economic Dynamics and Control*, 84, 43–57.

Magnus, J. R. and H. Neudecker (1980): "The Elimination Matrix: Some Lemmas and Applications," *SIAM Journal on Algebraic Discrete Methods*, 1, 422–449.

McCracken, M. and S. Ng (2015): "FRED-MD: A Monthly Database for Macroeconomic Research," .

Mertens, K. and M. O. Ravn (2013): "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States," 103, 1212–1247.

——— (2014): "A Reconciliation of SVAR and Narrative Estimates of Tax Multipliers," *Journal of Monetary Economics*, 68, S1 S19.

Milunovich, G. and M. Yang (2013): "On Identifying Structural VAR Models via ARCH Effects," *Journal of Time Series Econometrics*, 5, 117–131.

Montiel Olea, J. L., J. H. Stock, and M. W. Watson (2016): "Inference in Structural VARs with External Instruments," Manuscript.

Nakamura, E. and J. Steinsson (2018): "High Frequency Identification of Monetary Non-Neutrality: The Information Effect," *Quarterly Journal of Economics (forthcoming)*.

Normandin, M. and L. Phaneuf (2004): "Monetary Policy Shocks: Testing Identification Conditions under Time-Varying Conditional Volatility," *Journal of Monetary Economics*, 51, 1217–1243.

Pötscher, B. M. (1991): "Effects of Model Selection on Inference," *Econometric Theory*, 7, 163–185.

Primiceri, G. E. (2005): "Time Varying Structural Vector Autoregressions," *Review of Economic Studies*, 72, 821–852.

Ramey, V. A. (2011a): "Can Government Purchases Stimulate the Economy," *Journal of Economic Literature*, 49, 673–685.

——— (2011b): "Identifying Government Spending Shocks: It's All in the Timing," *Quarterly Journal of Economics*, 126, 1–50.

Ramey, V. A. and S. Zubairy (2018): "Government Spending Multipliers in Good Times and in Bad : Evidence from U.S. Historical Data," *Journal of Political Economy (forthcoming)*.

Rigobon, R. (2003): "Identification through Heteroskedasticity," *The Review of Economics and Statistics*, 85, 777–792.

Rigobon, R. and B. Sack (2003): "Measuring the Reaction of Monetary Policy to the Stock Market," *Quarterly Journal of Economics*, 118, 639–669.

Romer, C. D. and D. H. Romer (2010): "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," *American Economic Review*, 100, 763–801.

Sentana, E. and G. Fiorentini (2001): "Identification, Estimation and Testing of Conditionally Heteroskedastic Factor Models," *Journal of Econometrics*, 102, 143–164.

Shephard, N. (1996): "Statistical Aspects of ARCH and Stochastic Volatility," in *Time Series Models: in Econometrics, Finance and Other Fields*, ed. by D. Cox, D. Hinkley, and O. Barndorff-Nielsen, New York: Chapman & Hall, 1–68.

Shimizu, S., P. O. Hoyer, A. Hyvärinen, and A. Kerminen (2006): "A Linear Non-Gaussian Acyclic Model for Causal Discovery," *Journal of Machine Learning Research*, 7, 2003–2030.

Sims, C. (1980): "Macroeconomics and Reality," *Econometrica*, 48, 1–48.

Staiger, D. and J. H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65, 557–586.

Stock, J. H. (2008): "Recent Developments in Structural VAR Modeling," in *NBER Summer Institute: Methods Lectures*.

Stock, J. H. and M. W. Watson (2002): "Has the Business Cycle Changed and Why?" in *NBER Macroeconomics Annual*, ed. by M. Gertler and K. Rogoff, 159–230.

Trefethen, L. N. and D. Bau (1997): *Numerical Linear Algebra*, Philadelphia: Society for Industrial and Applied Mathematics.

Uhlig, H. (1997): "Bayesian Vector Autoregressions with Stochastic Volatility," *Econometrica*, 65, 59–73.

——— (2005): "What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics*, 52, 381–419.