

Identification Based on Higher Moments in Macroeconometrics

Daniel J. Lewis

Department of Economics, University College London, London, United Kingdom;
email: daniel.j.lewis@ucl.ac.uk

ANNUAL
REVIEWS **CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Annu. Rev. Econ. 2025. 17:665–93

First published as a Review in Advance on
June 5, 2025

The *Annual Review of Economics* is online at
economics.annualreviews.org

<https://doi.org/10.1146/annurev-economics-070124-051419>

Copyright © 2025 by the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See credit lines of images or other third-party material in this article for license information.

JEL codes: C30, C32, E50, E60



Keywords

identification, higher moments, structural shocks, SVAR, heteroskedasticity, non-Gaussianity

Abstract

In the last two decades, identification based on higher moments has attracted increasing theoretical attention and been widely adopted in empirical practice in macroeconometrics. This article reviews two parallel strands of the literature. The first is identification strategies based on heteroskedasticity of the structural shocks, which can provide additional covariance equations. The second exploits non-Gaussianity more generally of the structural shocks for identification, generally under the assumption of independence, based on the mature independent components analysis literature. I describe in detail the seminal identification results and discuss recent extensions. For each scheme, I describe parametric and nonparametric implementations and highlight prominent empirical applications. I also discuss key issues for the adoption of such strategies, including weak identification and the interpretability of statistically identified structural shocks. I further outline key areas of ongoing research, such as the blending of multiple sources of identifying information.

1. INTRODUCTION

Following Sims (1980), structural vector autoregressions (SVARs) have been the workhorse of causal inference in macroeconometrics. To recover dynamic causal effects or structural shocks from these models, the econometrician must identify a matrix of contemporaneous causal effects. Most traditional approaches to doing so are economic in nature. In particular, they either impose economically motivated restrictions on these contemporaneous impacts [e.g., short-run restrictions (Sims 1980), long-run restrictions (Blanchard & Quah 1989), sign restrictions (Uhlig 2005)] or supply external economic information to help pin down a unique set of structural parameters [e.g., external instruments (Stock 2008, Mertens & Ravn 2013)]. A separate strand of the literature that has gained popularity in recent years employs so-called statistical identification, in particular exploiting the information from higher moments. It is well-known that the second moments of reduced-form vector autoregression (VAR) innovations—the reduced-form covariance matrix—contain inadequate information to uniquely decompose the innovations into the structural shocks and their contemporaneous impacts. However, moments beyond the covariance can contain sufficient information to achieve identification under certain conditions. This insight dates to as early as 1928 (Wright 1928).

Two distinct, although closely connected, threads of this literature have developed. The first is based on heteroskedasticity, stemming from Sentana & Fiorentini (2001). If the variances of the structural shocks change through time, then there is not just a single reduced-form covariance matrix to exploit. Thus, an SVAR model may no longer be underidentified using the expanded set of second moments. Over the past two decades, this idea has been applied, extended, and generalized in multiple directions, most popularly to a small number of discrete variance regimes (Rigobon 2003) but also to smooth transition models (Lütkepohl & Netšunajev 2017) and parametric processes (Sentana & Fiorentini 2001, Lanne et al. 2010). More recently, Lewis (2021) synthesized earlier approaches in a distribution-free identification argument that accommodates all previously proposed processes and, indeed, essentially arbitrary persistent variance processes, including those with a state-space representation.

The second strand of the literature is based on the information offered by higher moments more generally and is often referred to by the shorthand of identification via non-Gaussianity of the structural shocks. Since Gaussian distributions are fully characterized by their first two moments, any identification scheme that successfully exploits higher moments must leverage some deviation of the structural shocks' distribution from Gaussianity. Due to the Darmois–Skitovich theorem (e.g., Darmois 1953, Skitovich 1953), if at least all but one of the structural shocks exhibit non-Gaussianity and the structural shocks are mutually independent, then the decomposition of innovations into structural shocks is unique (e.g., Comon 1994). This result is well-known and has been widely adopted in statistics and signal processing, as well as in economics more generally. Within macroeconometrics, it has been exploited in several different ways. These include nonparametric estimators based on independent component analysis (ICA) algorithms (Hyvärinen et al. 2010), maximum likelihood estimators (Lanne et al. 2017), pseudo-maximum likelihood (PML) estimators (Gouriéroux et al. 2017), and moment-based estimators (Guay 2021, Keweloh 2021, Lanne & Luoto 2021). Of course, heteroskedasticity of the structural shocks in general implies non-Gaussianity; for example, shocks coming randomly from two different variance regimes are unconditionally non-Gaussian. Thus, the two strands of the literature are intrinsically connected.

Statistical identification approaches are attractive because they avoid making potentially controversial economic assumptions about the underlying structural parameters. Indeed, the structural parameters are typically the objects of interest to the econometrician, so imposing

assumptions on them may be unappealing. Therefore, statistical identification, which can remain largely agnostic to these coefficients, is attractive when the econometrician is interested in testing hypotheses about structural parameters. However, these identification approaches do make assumptions of their own—on the statistical properties of the data. While the presence of heteroskedasticity in macroeconomic or financial data may be uncontroversial, to avoid weak identification (and, thus, for standard inference techniques to be reliable) the departures from homoskedasticity may need to be substantial in typical applications; the mere existence of heteroskedasticity is not enough in finite samples. If adequate departures from homoskedasticity or Gaussianity are not present, such models will be only weakly identified, as recognized by Magnusson & Mavroeidis (2014) and Nakamura & Steinsson (2018) and recently studied by Lee et al. (2022), Lewis (2022), and Montiel Olea et al. (2022), for example. While it is generally hard to test identification conditions related to higher moments, there has been recent progress in this direction (e.g., Guay 2021, Lütkepohl et al. 2021, Lewis 2022).

Another key assumption, at least in the non-Gaussian family of approaches, is that of mutual independence of the structural shocks. Mutual independence is much stronger than the usual assumption that the structural shocks are mutually and serially uncorrelated, since it rules out the possibility that their variances follow a common or factor process, for instance—a condition that seems tenuous in light of episodes like the Great Moderation. Several recent papers have aimed to relax the independence assumption to varying degrees (e.g., Guay 2021, Lanne & Luoto 2021, Mesters & Zwiernik 2024). A final challenge of statistical identification is labeling. Based on the statistical information contained in higher moments alone, models are only ever identified up to sign/scale and column permutations, or locally; there is no way to place an interpretable “label” on a structural shock, a column of impact coefficients, or an impulse response function (IRF). Some papers have sought to provide an initial labeling approach based on a purely statistical preference over column order, which guarantees global identification but no economic interpretation (e.g., Lanne et al. 2017). More typically, though, authors resort to economic information—meaning that these identification approaches are not completely devoid of economic assumptions.

This review describes the identification strategies named above in detail and outlines ongoing areas of research; the focus is macroeconometric settings, particularly SVARs. However, applications and similar theoretical contributions exist throughout economics. A leading example is errors-in-variables models, where the idea of exploiting higher moments for identification dates to at least 1950 (Reiersøl 1950; see, e.g., Gospodinov et al. 2017, Lewbel 1997). Heteroskedasticity has been used to identify simultaneous equation models more broadly (e.g., Klein & Vella 2010, Lewbel 2012, Prono 2014). Arguments described in this review for macroeconometric settings are very similar to those used to study income dynamics (e.g., Blundell et al. 2008, Bonhomme & Robin 2009). Higher moments have also been used in the literature on noninvertible models, where they offer novel identifying variation to distinguish between invertible and noninvertible models that are observationally equivalent up to second moments (e.g., Gospodinov & Ng 2015, Gouriéroux et al. 2019). Clearly, the applicability of the information contained in higher moments and in particular non-Gaussianity is broad in economics, and a complete treatment is beyond the scope of this review.

The remainder of this review proceeds as follows. Section 2 reviews the SVAR setting. Section 3 discusses identification approaches based on heteroskedasticity. Section 4 presents approaches based on non-Gaussianity. Section 5 describes the issues posed by weak identification as well as methods for robust inference and detection. Section 6 outlines key open questions in the literature and active research areas. Section 7 concludes by identifying several appealing avenues for ongoing research.

2. SETTING

Consider an $n \times 1$ vector of observed variables, Y_t , with an assumed mean of zero for simplicity (for a comprehensive treatment of SVARs, see Kilian & Lütkepohl 2017). A standard SVAR has the form

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + u_t, u_t = B \epsilon_t, \quad 1.$$

or, more compactly,

$$A(L)Y_t = B \epsilon_t. \quad 2.$$

Here, u_t are the reduced-form SVAR innovations and ϵ_t are mean-zero mutually orthogonal and serially uncorrelated structural shocks, or

$$E[\epsilon_t] = 0, E[\epsilon_{it}\epsilon_{js}] = 0, \forall i \neq j, s \neq t, E[\epsilon_t \epsilon_t'] = \Lambda, \quad 3.$$

where Λ is diagonal. Crucially, B is the $n \times n$ matrix of contemporaneous causal effects, assumed to be invertible.¹ The coefficients of the lag polynomial, $A(L)$, are consistently estimable via ordinary least squares (OLS). The SVAR identification problem pertains to B . In particular, the covariance of the reduced-form innovations, u_t ,

$$\Sigma = E[u_t u_t'], \quad 4.$$

provides only $n(n+1)/2$ unique equations, but there are n^2 unknown parameters to be identified in the $n \times n$ matrix B and the n -entry diagonal covariance of ϵ_t , even after a suitable scale normalization, giving n restrictions, is imposed (via either a unit diagonal for B or identity covariance for ϵ_t).² Indeed, the impact coefficients of B are identified only up to orthonormal rotations. To understand this point, impose the unit variance normalization, so that $\Lambda = I_n$. Let $Q \in \mathcal{O}^n$, where $\mathcal{O}^n$ is the space of all $n \times n$ orthonormal matrices. Then,

$$\Sigma = B I_n B' = B B' = (BQ) Q' I_n Q (Q' B') = B^* B^{*'}, \quad 5.$$

where $B^* = BQ$. In other words, B and B^* are observationally equivalent up to second moments of u_t without further restrictions.

As discussed in Section 1, there are now two options to identify B uniquely. Either restrictions must be imposed on B , to reduce the number of free parameters from n^2 toward the $n(n+1)/2$ equations available in Σ (or, in the case of sign restrictions, reduce the space of permissible rotations from $\mathcal{O}^n$ to some subset of $\mathcal{O}^n$), or additional information must be furnished, through either external variables or further moments of u_t . This review focuses on the latter option, exploiting higher moments of the data. Two notions of identification are considered: Local identification means that there is some neighborhood around B in which no other parameters are observationally equivalent, while global identification means that there exist no other observationally equivalent parameters. Below, higher moments alone only ever identify B up to scale and column permutations, so they are at most locally identifying without economic or other restrictions.

¹Note that the invertibility assumption is not innocuous and is studied by a growing literature (for a discussion of such issues, see Fernández-Villaverde et al. 2007). Gouriéroux et al. (2019) consider the interaction of non-Gaussianity and noninvertibility.

²Note that the identity covariance normalization results in a sign indeterminacy for the columns of B , or, equivalently, the structural shocks, that is not associated with the unit diagonal normalization. That translates into local identification below, resulting in solutions that are unique only up to sign in addition to column permutations.

3. HETEROSKEDASTICITY

The intuition for how identification via heteroskedasticity can solve the SVAR identification problem is straightforward. Observational equivalence is based on the single set of equations, contained in Σ , available under homoskedasticity. However, if there are multiple values for the variances of the structural shocks through time, then the number of available equations scales linearly. The key assumption permitting identification is that B remains constant, even as the variances change. Then, the number of new free parameters added for each variance regime is only n , fewer than the $n(n+1)/2$ equations added, provided that those equations are linearly independent.

3.1. Identification from Variance Paths

Sentana & Fiorentini (2001) provide the first formal result exploiting heteroskedasticity directly for identification.³ Let Λ_t denote the diagonal covariance matrix of the structural shocks at time t , which follows an arbitrary stochastic process. Then, their proposition 3 shows that if the paths of the n diagonal elements, λ_{it} , through time have full rank, then B is uniquely determined up to permutations (and scale/sign normalization). The condition allows at most one shock to be homoskedastic. Indeed, the requirement of $n-1$ dimensions of linearly independent time-varying volatility is shared by all schemes exploiting heteroskedasticity below, and mirrored by the requirement of at least $n-1$ non-Gaussian shocks in the next section. In general, this result provides substantial overidentification, as discussed below. However, a major limitation of this result for practical use is that identification is based on the time path of Λ_t , and thus the time path of Σ_t , denoting the time-specific covariance of the reduced-form innovations. In general, these time-specific reduced-form covariances are not consistently estimable without further particular parametric assumptions on the variance process for the structural shocks: As T increases, the sample size informative for time t stays fixed without any further restrictions. Sentana & Fiorentini's (2001) solution is to impose a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) functional form for the structural variance process, under which the variances are deterministic functions of past data and consistently estimable parameters. A relatively small empirical literature has employed the Sentana & Fiorentini (2001) approach directly, based on estimating a GARCH process for the structural shocks by (pseudo-)maximum likelihood, and thus estimating the structural parameters of interest (e.g., King et al. 1994, Normandin & Phaneuf 2004, Bouakez & Normandin 2010, Lütkepohl & Milunovich 2016).

3.2. Variance Regimes

Rigobon (2003) makes perhaps the best-known, and most widely applied, contribution to this literature. In a special case of Sentana & Fiorentini's (2001) results, he argues that if there are two discrete regimes for the structural variances, then B is uniquely determined provided that Λ_1 and Λ_2 are not scalar multiples of each other. Simple equation counting shows why: $2 \times n(n+1)/2 = n^2 + n$, which is the number of structural parameters in B and two diagonal variance matrices, after n elements are normalized for scale. If there are more than two regimes, or values for the structural variances, the model is overidentified. Rigobon's proposition 1 gives rise to the well-known identification condition in the two-regime case: Identification holds as long as none of the variances change proportionally across regimes. Generalized for $n \geq 2$, this condition reads as

$$\lambda_{2i}/\lambda_{1i} \neq \lambda_{2j}/\lambda_{1j}, \forall i \neq j. \quad 6.$$

³However, the result appears earlier in an unpublished working paper (Sentana 1992).

There are extensions allowing for additional common shocks (requiring further regimes) (Rigobon 2003, proposition 2). Crucially, the econometrician is not required to know the precise dates of the two variance regimes or, indeed, that the variances follow a discrete process at all. Provided that the econometrician specifies the regimes such that the associated covariance estimators are consistent for distinct variance matrices (without proportional entries), the identification condition is met (Rigobon 2003, proposition 3). However, the better specified, and thus more distinct, the regimes are, the stronger identification is likely to be. Additional regimes that provide overidentifying information allow the econometrician to test the modeling assumptions via the overidentifying restrictions, such as the assumption that B stays fixed over time.

In the special case of two regimes, where the model is just-identified, a convenient closed-form solution exists. As noted by Lanne et al. (2010), for example, B is identified as the left eigenvectors of the matrix $\Sigma_2 \Sigma_1^{-1}$:

$$\Sigma_2 \Sigma_1^{-1} = B \Lambda_2 \Lambda_1^{-1} B^{-1}, \quad 7.$$

where $\Lambda_2 \Lambda_1^{-1}$ is diagonal, containing the eigenvalues, and thus B contains the eigenvectors. These eigenvectors are unique (up to normalization and order) provided that there are no repeated eigenvalues—no variances change proportionally across regimes. This result also clarifies how partial identification can occur: If there are repeated eigenvalues, then the columns of B (eigenvectors) corresponding to the nonrepeated entries are still identified, but only the column space associated with repeated values is identified. The case of partial identification has received relatively little formal attention, although Lütkepohl et al. (2020) establish asymptotic theory under partial identification (along with tests to determine the extent to which point identification fails) and Bacchiocchi et al. (2024) derive identified sets resulting from partially identifying heteroskedasticity and sign or zero restrictions. In general, estimation of these regime-based models proceeds via the generalized method of moments (GMM). Moments take the form

$$\phi(B, \Lambda_1, \dots, \Lambda_N, \eta_t) = \begin{pmatrix} 1[t \in R_1] (E[\eta_t \eta_t'] - B \Lambda_1 B') \\ \vdots \\ 1[t \in R_N] (E[\eta_t \eta_t'] - B \Lambda_N B') \end{pmatrix}, \quad 8.$$

where $1[t \in R_r]$ is an indicator for whether observation t belongs to the r th regime. However, maximum likelihood approaches are also available, and Brunnermeier et al. (2021), Bacchiocchi & Kitagawa (2023), and Bacchiocchi et al. (2024), for example, adopt fully Bayesian frameworks.

Under additional assumptions, such that the variance of only one shock changes through time with the other(s) remaining homoskedastic, Rigobon & Sack (2004) show that an instrumental variables (IV)-type estimator is available for a typical parameter of interest. In particular, consider $n = 2$ and let B_{ij} be the coefficient of interest, and assume that it measures the effect of the shock whose variance does change, say j , on some variable, say i . In this case, under the unit diagonal normalization,

$$\frac{\Sigma_{2,ij} - \Sigma_{1,ij}}{\Sigma_{2,jj} - \Sigma_{1,jj}} = \frac{B_{ij}\lambda_{2j} + B_{ji}\lambda_i - B_{ij}\lambda_{1j} - B_{ji}\lambda_i}{\lambda_{2j} + B_{ji}^2\lambda_i - \lambda_{1j} - B_{ji}^2\lambda_i} = B_{ij} \frac{\lambda_{2j} - \lambda_{1j}}{\lambda_{2j} - \lambda_{1j}} = B_{ij}. \quad 9.$$

The left-hand side is equivalent to

$$\frac{E[u_{it}Z_t]}{E[u_{jt}Z_t]}, \quad 10.$$

where the instrument Z_t is given by

$$Z_t = \frac{T}{T_2} 1[t \in R_2] u_{jt} - \frac{T}{T_1} 1[t \in R_1] u_{jt}, \quad 11.$$

and T_r denotes the number of observations in regime r .

For the estimators described above, provided that identification is strong—that is, Equation 6 is satisfied—inference can proceed using the standard asymptotic results associated with each estimation strategy. However, the final formulation of the identification strategy (under the additional assumption that one shock is homoskedastic), Equation 10, presents an analogy to IV estimation, which makes clear that weak identification is possible, in this case when the variance of shock j changes little between the two regimes (for further discussion, see Section 5).

A key distinction between the Sentana & Fiorentini (2001) and Rigobon (2003) results is the type of heteroskedasticity accommodated. In principle, the former applies to both unconditional and conditional heteroskedasticity, although Sentana & Fiorentini (2001) focus on conditional heteroskedasticity. The latter exploits unconditional heteroskedasticity, and it remains to determine the variance regimes. In some cases, external information proves helpful; perhaps the most popular example is, in daily financial data, to use dates corresponding to announcements, especially monetary policy announcements, as a high-variance regime, and dates far away from such announcements as a control regime (e.g., Rigobon & Sack 2004, Nakamura & Steinsson 2018). Applying the same logic to continuous time periods like the Great Moderation, for instance, is more problematic, since doing so may mask considerable variation (periods of larger shocks within an otherwise low-variance interval). Rigobon (2003) uses narrative information on the dates of tranquil and crisis periods in Latin American debt markets to define regimes for identification. Alternatively, Rigobon & Sack (2003) propose to estimate the regime dates based on realized volatility, in particular by comparing rolling averages of squared reduced-form SVAR innovations with the average levels of the squared innovations to determine high- and low-variance periods. In population, the precise regime breaks are not crucial, but in finite samples, estimation error can lead to more muted variance changes and, thus, weak identification.

An additional drawback of estimating regimes, however, that has not been explored in the literature is that the realizations of squared innovations used to determine variance regimes are of course driven by realized values of the structural shocks. Depending on B , it might be the case that, for instance, large positive realizations of two shocks lead to larger squared innovations than realizations of the same magnitude but opposite signs. Thus, conditional on the realized value of the squared innovations, the shocks may not be orthogonal. As a result, it need not be the case that the covariance matrix of the structural shocks is diagonal conditional on estimated regime membership, introducing bias.

When regimes are unknown, perhaps a more natural option is to estimate them parametrically, using a Markov switching model. This is precisely the innovation of Lanne et al. (2010). In a straightforward extension, they propose a maximum likelihood estimator exploiting the Rigobon (2003) identification result; the identification conditions remain unchanged. Herwartz & Lütkepohl (2014) combine a Markov switching model with more conventional short- and long-run restrictions. Lütkepohl & Woźniak (2020) develop a Bayesian implementation of this approach and use it to test overidentifying restrictions on B .

Another important extension of the regime-based model is the so-called smooth transition approach, in which the covariances are assumed to move as a convex combination between two matrices in a continuous manner. Lütkepohl & Netšunajev (2017) propose such a model. Concretely, the reduced-form covariance, Ω_t , follows the law of motion:

$$\Omega_t = [1 - F(s_t)]\Sigma_1 + F(s_t)\Sigma_2, \quad 12.$$

where $F(\cdot)$ is a parametric function. In particular, they propose

$$F(\gamma, b, s_t) = (1 + \exp[-\exp(\gamma)(s_t - b)])^{-1}, \quad 13.$$

where γ is a slope parameter determining the speed of transition, b is a location parameter, and s_t is the so-called transition variable that governs the state at time t . It can be either a random variable (lagged inflation in Lütkepohl & Netšunajev 2017) or a deterministic variable, like t . If the former, it must be exogenous in order for standard inference results to hold. As $\gamma \rightarrow \infty$, the model approaches a threshold model, where $\Omega_t = \Sigma_1$ for $s_t < b$ and $\Omega_t = \Sigma_2$ for $s_t > b$. As $\gamma \rightarrow 0$, the model becomes unidentified (variances are constant), so weak identification is a concern for small values of γ . Conditional on a set of parameters in the transition equation, (γ, b) , identification follows by Rigobon's (2003) argument, applied to Σ_1 and Σ_2 . In practice, Lütkepohl & Netšunajev (2017) note that (γ, b) must be identified and that the choice of s_t may be important. For example, if s_t does not evolve with the underlying structural variances, then identification will likely be weak or nonexistent. The authors propose to estimate the model via Gaussian maximum likelihood across a grid of parameters for (γ, b) using a two-step procedure, alternating between estimating the reduced-form and structural parameters. Then, the final log likelihoods can be compared across the grid for (γ, b) .

3.3. Unconditional Moments

All of the papers described in the preceding section adopt different forms for the variance process but rely on the same key insight: Multiple values for the structural variances offer additional covariance matrices from which B may be identified. These approaches can all be thought of as relying on the paths of the structural or reduced-form variances for identification. A much smaller literature argues for identification based on the properties of the process governing the evolution of the variances through time. Milunovich & Yang (2013) revisit identification based on ARCH-type functional forms. They formulate the mapping between parameters of the reduced-form GARCH process for the innovations and those of the structural GARCH process for the shocks. They assume that the structural ARCH coefficient matrix is diagonal. Based on the Jacobian of these equations, they show that B and the structural GARCH parameters are jointly locally identified from the reduced-form GARCH parameters provided that there is at most one zero on the diagonal of the structural ARCH coefficient matrix. This condition amounts to a requirement that at most one shock is homoskedastic, mirroring the arguments presented above.

Lewis (2021) takes a different approach to argue that B is identified from the unconditional autocovariances of the squares of the reduced-form SVAR innovations. Under the slightly stronger assumption that the shocks are a martingale difference sequence with respect to past shocks and current and past volatilities, and finite fourth moments, the following equations hold:

$$\text{cov}[\text{vec}(u_t u_t'), \text{vec}(u_{t-s} u_{t-s}')] = (B \otimes B) G \text{cov}[\lambda_t, \text{vec}(\epsilon_{t-s} \epsilon_{t-s}')](B \otimes B)', s > 0, \quad 14.$$

$$E[u_t u_t'] = BE[\Lambda_t]B'. \quad 15.$$

Here, G is a selection matrix of zeroes and ones. The main result is that this system of equations has a unique solution for B , $\text{cov}[\lambda_t, \text{vec}(\epsilon_{t-s} \epsilon_{t-s}')]$, and $E[\Lambda_t]$, up to normalization and column order, provided that, for some lag s , the $n \times (n^2 + 1)$ matrix $\begin{bmatrix} \text{cov}[\lambda_t, \text{vec}(\epsilon_{t-s} \epsilon_{t-s}')'] & E[\lambda_t] \end{bmatrix}$ has a rank of at least two and no proportional rows. If there are no ARCH effects, then $\text{cov}[\lambda_t, \text{vec}(\epsilon_{t-s} \epsilon_{t-s}')'] = \text{cov}(\lambda_t, \lambda_{t-s}')G'$, the autocovariance matrix of λ_t at lag s . Applied to $\text{cov}[\lambda_t, \text{vec}(\epsilon_{t-s} \epsilon_{t-s}')']$, the rank condition essentially means that the autocovariance structures of the shock volatilities are not proportional. This rank condition will be satisfied if the volatility processes have persistence

coming from a source other than a lower-dimensional factor structure, for example. Augmented with $E[\lambda_t]$, the condition states that even if the autocovariance structures are proportional, identification will still hold as long as the constants of proportionality are not equal to the ratios between the mean structural variances. Notably, this condition can be interpreted as allowing at most one homoskedastic shock, much like the earlier schemes. If the identification conditions for any of the preceding schemes hold, then the rank condition above is satisfied. In this sense, the Lewis (2021) argument nests all previous identification schemes based on heteroskedasticity, without relying on parametric features for identification or to ensure consistent estimability of identifying moments.

The intuition behind the argument is that if the shocks are serially uncorrelated (and, in fact, are martingale difference sequences), then the only persistence in the squared innovations is that of the variance process (and/or ARCH effects). Then B , or rather $(B \otimes B)$, contains the coefficients relating the autocovariance of the squared reduced-form innovations to the autocovariance of the squared structural shocks.⁴

Lewis's (2021) argument has several important features. First, it is distribution-free, so it can be implemented without assuming any form for the identifying heteroskedasticity, for instance, by simply estimating Equations 14 and 15 by GMM. Second, given standard macroeconomic sample sizes, it will often be challenging, however, to estimate the required higher moments precisely, so in practice it may be more appealing to view it as a general-purpose argument that shows that a very wide range of parametric models will identify the structural parameters of interest. Provided that a parametric model satisfies the rank condition, then B will be identified from the associated likelihood, if it is well-specified. Finally, distinct from any previous arguments based on heteroskedasticity, it permits identification based on volatility models that include state variables. For example, the autoregressive log stochastic volatility model is very popular in empirical research (e.g., Cogley & Sargent 2005, Primiceri 2005) but previously was not compatible with any available identification arguments, since they all require consistent estimation of the path of (reduced-form) volatilities (e.g., Bertsche & Braun 2022, developed concurrently). Since the Lewis (2021) argument instead requires consistent estimation of unconditional moments of the variances, it is compatible with state-space volatility models.

Although this review focuses on identification as opposed to estimation and inference, one common estimation issue across these identification schemes warrants further discussion. Typically, SVARs are estimated in a two-step process. In the first step, the reduced-form VAR is estimated via OLS, and in the second step, the implied estimated innovations are treated as data to estimate the structural parameters. For estimation purposes, these two steps are often treated as entirely separable (although inference may adjust for estimation error in $\hat{u}_t$). However, identification via heteroskedasticity motivates generalized least squares-type estimators for the entire model, since they may offer efficiency gains if heteroskedasticity is present. In practice, maximum likelihood estimation alternates among estimating the reduced-form parameters and structural parameters, computing time-varying variances and thus weights, and then updating, until convergence (as described above for, e.g., Lütkepohl & Netšunajev 2017). Moreover, without additional assumptions, inference on the structural parameters (particularly IRFs) is complicated. Indeed, the estimation error in $\hat{u}_t$ cannot be ignored for the purpose of inference on $\hat{B}$ and IRFs in general. However, if the shocks, ϵ_t , are assumed to follow a symmetric distribution, one can show

⁴The structure of the argument, particularly in the simplified motivating example presented in this review, is not dissimilar to the seminal argument of Blundell et al. (2008), who identified coefficients of an income process in panel data.

that the estimators of the reduced-form covariances and those of the reduced-form VAR parameters are asymptotically independent, which means that estimation error in $\hat{u}_t$ does not necessitate adjustments in estimating the asymptotic variance of $\hat{B}$ (for further discussion, see especially theorem 2.1 of Brüggemann et al. 2016 and footnotes 15 and 20 of Lewis 2021; for a detailed discussion of estimation for many of the schemes described above, see Kilian & Lütkepohl 2017).

4. NON-GAUSSIANITY

While the previous section considers the use of a particular type of higher moment for identification, identification based on non-Gaussianity (and some independence-like assumption for the structural shocks) exploits moments beyond the second generically. Some approaches take a stand on what types of deviation from Gaussianity are likely to be informative in macroeconomic data, while others are more flexible. Note that, in general, heteroskedasticity generates non-Gaussianity in the structural shocks, even if the underlying (standardized) disturbances are themselves Gaussian, so identification via heteroskedasticity can be thought of as a special case of identification based on non-Gaussianity.

This approach can be motivated by the idea that structural shocks should be independent (or, at least, more than uncorrelated), as discussed by Keweloh (2024), for instance. For example, the shocks $\epsilon_{1t} \sim \mathcal{N}(0, 1)$ and $\epsilon_{2t} = \epsilon_{1t}^2 - 1$ are uncorrelated but exhibit dependence that is inconsistent with how most macroeconomists think about structural shocks. Independence assumptions rule out these types of dependence. Given such an assumption, non-Gaussianity can identify the shocks. Consider the moment $E[\epsilon_{1t}^2 \epsilon_{2t}]$, a “coskewness” (i.e., skewness in the product of two shocks) condition that measures the dependence of the two shocks. If $E[\epsilon_{1t}^2 \epsilon_{2t}] \neq 0$, then the first shock’s size is informative for the sign of the second shock, and independence is violated; the key is to find an orthogonal rotation of the structural shocks such that independence holds. If the shocks are Gaussian, the moment is mechanically zero, but if they are non-Gaussian, such moments are informative for B .

4.1. The Main Result

The idea that non-Gaussianity in general suffices to identify B follows from the Darmois–Skitovich theorem (Darmois 1953, Skitovich 1953), a generalization of the earlier Kac–Bernstein theorem (Kac 1939, Bernstein 1941). This result states that if the elements of $\zeta = (\zeta_1, \dots, \zeta_n)'$ are independent random variables and $\mu'\zeta$ and $\beta'\zeta$ are independent (for nonzero μ and β), then all ζ_i are Gaussian. As a consequence, Comon (1994, theorem 11 and corollary 13) shows that a decomposition of the form $u_t = B\epsilon_t$ is unique up to column order and scale provided that at most one component of ϵ_t is Gaussian and the components of ϵ_t are independent. This follows from the use of a contrast function, $\Psi(\cdot)$. According to Comon (1994, definition 5), for some random vector e_t , $\Psi(f_{e_t}) = \Psi(f_{\epsilon_t})$, where f_{e_t} denotes the density of e_t , if and only if e_t is a scaled permutation of ϵ_t . Furthermore, if ϵ_t are independent, then $\Psi(f_{M\epsilon_t}) \leq \Psi(f_{\epsilon_t})$, where M is invertible, so maximizing the contrast function will return (a scaled permutation of) ϵ_t . Intuitively, $\Psi(\cdot)$ measures the (negative) deviation of a random vector from independence, and any candidate shock vector e_t cannot be any more independent than the original shocks, ϵ_t .

It remains to propose candidate functions with these properties. Early results on identifiability under non-Gaussianity are presented by Geary (1941) and Reiersøl (1950). A notable feature of all of these results is the independence assumption on ϵ_t , which is notably stronger than the mutually orthogonal and serial uncorrelatedness assumption typically employed elsewhere in the SVAR literature. I return to this topic in Section 4.4.

4.2. Independent Components Analysis

Results like Comon's (1994) are central to the ICA literature, which originated the idea of identification via non-Gaussianity in modern macroeconometrics. ICA focuses on deconvolutions of the form $u_t = B\epsilon_t$, where ϵ_t are independent, and is a key technique in the signal processing and neural network literatures. While there are many implementations of ICA, the basic idea is, for an initial set of candidate standardized shocks (e.g., those arising from a Cholesky factorization, say $\tilde{\epsilon}_t$), to find the orthonormal rotation Q for which the shocks $e_t = \tilde{\epsilon}_t Q$ are as independent as possible, in practice by maximizing some contrast function. For example, Comon (1994) proposes to use a feasible approximation to the (negative of) the Kullback–Leibler divergence of the joint density of e_t from the density under the independence of $e_{1t}, \dots, e_{nt}$ (i.e., mutual information). This approximation can be expressed in terms of the cumulants of u_t and Q .

I first consider nonparametric ICA. The ICA literature is vast, and a comprehensive discussion is outside of the scope of this review (for a review of this literature, see, e.g., Hyvärinen 2013, Mesters & Zwiernik 2024). I focus on contributions that have directly affected macroeconometrics. One such variant is the FastICA algorithm (Hyvärinen 1999), which is in wide use and available in many statistical software packages (many alternatives exist; see, e.g., the JADE algorithm of Cardoso & Souloumiac 1993). FastICA exploits results showing that the negentropy (which measures the distance between some distribution and the normal distribution) can be better approximated by the maximum entropy principle than by cumulants (e.g., Hyvärinen 1997) and, furthermore, that minimizing the mutual information is approximately equivalent to maximizing the negentropy, or the degree of non-Gaussianity, of each shock. In practice, the algorithm uses an approximation to the negentropy of $e_{it} = c_i' u_t$, $J_b(\cdot)$,

$$J_b(c_i) = (E[b(c_i' u_t)] - E[b(z)])^2, \quad 16.$$

for some $n \times 1$ vector c_i and (nonquadratic) function $b(\cdot)$, where z is standard normal. For symmetric variables, this is a generalization of the cumulant-based approximation presented by Comon (1994), where $b(e_{it}) = e_{it}^4$. Then, the contrast function is $\sum_{i=1}^n J_b(c_i)$, which is maximized over $C = B^{-1}$ (where the i th row of C is c_i). Hyvärinen (1999) proposes a number of different choices for $b(a)$ implemented in the algorithm, including $1/k_1 \log \cosh(k_1 a)$, $-1/k_2 \exp(-k_2 a^2/2)$, and $1/4 a^4$, motivated by efficiency and robustness considerations. The FastICA algorithm then proceeds to maximize the negentropy given $b(\cdot)$ using a computationally efficient fixed point algorithm. While most of the ICA literature assumes independent and identically distributed shocks, a concern that I discuss in detail in Section 4.4, some recent algorithms have been extended to accommodate heteroskedasticity (e.g., gJADE; Matilainen et al. 2015).

Results on the statistical properties of estimators are relatively rare in the ICA literature. In particular, consistency results for the estimators resulting from many algorithms are hard to find, and some are in fact inconsistent (for discussion, see Gouriéroux et al. 2017). The consistency and asymptotic normality of the FastICA algorithm were established by Wei (2015), although expressions for the asymptotic variance (assuming consistency) originated with Shimizu et al. (2006), for example. Bonhomme & Robin (2009) provide earlier consistency and asymptotic normality results for a modification of the JADE algorithm.

Hyvärinen et al. (2010) were likely the first to exploit modern non-Gaussianity results to identify an SVAR, using financial data. They consider an SVAR of four global stock indices and estimate the reduced form before applying the FastICA-based LiNGAM algorithm of Shimizu et al. (2006). They impose a sparsity penalty in the estimation to find that more than half of the impact coefficients are zero and that $\hat{B}$ can be permuted to a lower-triangular matrix, with strong spillovers between certain indices. Moneta et al. (2013) employ the same methodology but in an economics outlet, with applications to both firm growth and the effects of monetary policy.

4.3. Likelihood Approaches

The next generation of results to identify SVARs using non-Gaussianity rely on likelihood approaches instead of nonparametric ICA. Lanne et al. (2017) assume that the shocks are sequences of independent and identically distributed and mutually independent disturbances with variances λ_i , with at most one shock Gaussian. Their identification result is based on those above. They propose a maximum likelihood estimator in terms of the density $f_{i,\lambda_i}(x, \theta_i) = \lambda_i^{-1/2} f_i(x/\lambda_i^{1/2}, \theta_i)$, but in practice assume that each shock follows a Student t distribution. Standard asymptotic properties for maximum likelihood estimation hold, provided that the choice of densities is correct—a rather heroic assumption, given that there are infinitely many ways to model non-Gaussianity, and presenting a challenge to empirical users. They propose a computationally simpler three-step estimator that is efficient if the shocks follow symmetric distributions. In an extension, they are able to relax the assumption of temporal independence to no serial correlation, which admits time-dependent heteroskedasticity for the shocks, provided that the volatility processes remain independent. They also provide an approach for refining local identification to global identification, which is discussed in detail in Section 6.4. In an empirical application, they study the relationship between the macroeconomy and financial conditions and are able to marginally reject a conventional recursive structure in their SVAR.

Contemporaneously, Gouriéroux et al. (2017) instead consider PML estimation. They assume that the shocks ϵ_t are independent, with at most one Gaussian. Importantly, they establish consistency for the PML estimator even when the likelihood is misspecified, provided that the misspecified model is identified. Additional results offer testable implications of Gaussianity of two or more shocks for the observed data, Y_t . They also provide expressions for the asymptotic variance. In an empirical application involving real activity, inflation, the Fed Funds rate, and oil prices, they are able to reject two recursive schemes, depending on the chosen real activity variable. Additional PML results are presented by Hyvärinen et al. (2001).

More recently, Jarociński (2024) proposed to estimate the effects of four different dimensions of monetary policy based on non-Gaussianity using maximum likelihood estimators. His baseline approach uses the Student t distribution, as done by Lanne et al. (2017), but he also proposes an alternative estimator allowing for dependence. In particular, he allows for endogenously determined dependence in the tails of the shock distributions by designing a new partially dependent multivariate t distribution that nests both independent and multivariate t distributions as extreme cases. In the data, the level of dependence is found to be small, leading to minimal changes in results, but this contribution may be of interest in many applications.

Elsewhere, Chen & Bickel (2006) avoid the choice of likelihood by proposing a semiparametric estimator based on the efficient score function, which they show is asymptotically efficient. Hafner et al. (2023) propose to maximize a kernel estimate of the likelihood in another nonparametric approach. Fiorentini & Sentana (2023) propose mixtures of normals PML estimators. They prove that, while for other likelihood approaches estimates of impact coefficients (and autoregressive parameters) will generally be consistent under misspecification, they are inconsistent for VAR intercepts and shock moments. Their proposed estimators provide consistent estimates of all parameters. Maxand (2020) considers maximum likelihood estimation of models with possibly more than one Gaussian shock, and provides tests for the number of such shocks. Bayesian implementations of identification based on non-Gaussianity can be found in papers by Lanne & Luoto (2020) (t -distributed shocks to assess the plausibility of sign restrictions holding), Anttonen et al. (2023) (generalized skewed t -distributed shocks; Markov-chain Monte Carlo methods are proposed), Braun (2021) (Dirichlet process mixture model), Keweloh et al. (2023b)

(skewed t -distributed shocks and potentially invalid proxy variables), and Lanne et al. (2023b) (t -distributed errors and a GARCH process for the shock volatilities).

4.4. Moment-Based Approaches

More recently, the literature has turned to moment-based estimators exploiting non-Gaussianity. Typically, these results relax the independence assumption and instead require uncorrelated shocks that additionally satisfy either a number of zero restrictions on coskewness, that is, third moments that are a function of two distinct shocks,

$$E[\epsilon_{it}^2 \epsilon_{jt}] = 0, \quad i \neq j, \quad 17.$$

or fourth moments that are a function of two distinct shocks,

$$E[\epsilon_{it}^2 \epsilon_{jt}^2] - 1 = 0, \quad E[\epsilon_{it}^3 \epsilon_{jt}] = 0, \quad i \neq j. \quad 18.$$

These coskewness and cokurtosis conditions are implied by the stronger (and previously maintained) independence assumption. Unsurprisingly, this exercise (like similar efforts to allow dependence in the likelihoods above) presents a trade-off, since additional restrictions implied by independence can improve identification when valid. By virtue of selecting specific moments, all such approaches rely on particular deviations from Gaussianity, but in practice, the identification condition is heuristically thought of as “at least $n - 1$ non-Gaussian shocks.”

One set of identification results focuses on the use of fourth moments, particularly excess cokurtosis restrictions. These results date to Bonhomme & Robin (2009), who show that B is identified (up to sign and permutations) from the covariance and cokurtosis (e.g., the collection of equations $E[u_{it}u_{jt}u_{lt}u_{mt}]$) of reduced-form innovations u_t , provided that the cokurtosis tensor of ϵ_t is diagonal, as implied by independence, and that at most one shock has zero excess kurtosis. Their proof is based on spectral arguments, in a factor model setting. They establish consistency and asymptotic normality of GMM estimators, as do the following authors. Guay (2021) considers the same moment conditions as Bonhomme & Robin (2009) but establishes local identification under the assumption of zero excess cokurtosis by studying the rank of the Jacobian, which facilitates additional results discussed below. Keweloh (2021) works from slightly different moments, written instead in terms of the underlying shocks, which, under a mutual independence assumption and normalization, have the form

$$E[\epsilon_{it}^2] = 1, \quad 19.$$

$$E[\epsilon_{it}\epsilon_{jt}] = 0, \quad i \neq j, \quad 20.$$

$$E[\epsilon_{it}^2 \epsilon_{jt}^2] = 1, \quad i \neq j, \quad 21.$$

$$E[\epsilon_{it}^3 \epsilon_{jt}] = 0, \quad i \neq j. \quad 22.$$

After the set of permissible B matrices is restricted to those matching the permutation/scaling rule of Lanne et al. (2017) and under the assumption that $n - 1$ shocks exhibit excess kurtosis, the covariance and cokurtosis conditions globally identify B , following an argument similar to that of Comon (1994). Note that while each author may write the assumptions and results somewhat differently, these identification conditions are indeed the same, exploiting all third and/or fourth moment restrictions implied by independence. The only difference is that writing the moments in terms of ϵ_t instead of u_t entails fewer moments and unknowns but requires the inversion of B , both of which may have finite sample consequences.

Lanne & Luoto (2021) argue that B is identified if $E[\epsilon_{it}^3 \epsilon_{jt}] = 0$ for at least $n(n - 1)/2$ combinations of $i \neq j$. This condition appears weaker than those in the preceding three papers

(Bonhomme & Robin 2009, Guay 2021, and Keweloh 2021). However, this result does not hold under the stated assumptions; rather, it requires all cokurtosis restrictions (symmetric, i.e., Equation 21, and asymmetric, i.e., Equation 22) implied by independence to be satisfied—the same as the previous papers. Lanne & Luoto (2021) rely on these conditions in their proof, which establishes local identification via the Jacobian of the identifying moments. Keweloh (2021) shows by counterexample that under the stated conditions, the model is not identified up to sign and permutation. Ultimately, all of the above studies require exactly the same restrictions on the dependence of the shocks. Lanne et al. (2023a) subsequently introduce the additional assumption that $n - 1$ shocks have excess kurtosis of the same sign, under which they achieve identification using only $n(n - 1)/2$ higher moments—all of the symmetric cokurtosis conditions implied by independence. However, asymptotic normality also requires $n(n - 1)/2$ asymmetric cokurtosis restrictions. They argue that this number is potentially far lower than the moments required by Keweloh (2021). Nevertheless, they still require all fourth moments implied by independence to hold for their asymptotic theory, even if they are not exploited for identification. They propose a moment selection procedure, although it is unlikely to perform well in realistic sample sizes in the larger SVARs, where it would be most useful. Karamysheva & Skrobotov (2022) extend the identification results of Lanne & Luoto (2021) to the so-called AB model, where $A_0 u_t = B \epsilon_t$ and A_0 must also be identified.

Another strand of this literature uses third moments, or skewness, for identification, either separately or in conjunction with kurtosis. Again, all restrictions on coskewness implied by independence are used for identification. In parallel to their result using kurtosis, Bonhomme & Robin (2009) show that under the restriction of zero coskewness (i.e., $E[\epsilon_{it}^2 \epsilon_{jt}] = 0$, $i \neq j$), provided that at most one shock has zero skewness, B is identified, following similar spectral arguments. Guay's (2021) identification argument also applies to sets of moments including skewness conditions alongside kurtosis conditions. The identification condition holds when all but one shock exhibits skewness and/or nonzero excess kurtosis. Moreover, Guay establishes partial identification results, such that the parameters corresponding to all skewed and/or nonmesokurtic shocks are identified regardless of the properties of the remaining shocks. He further shows that the rank can be decomposed into the sum of the ranks of certain blocks of the Jacobian. Importantly, this decomposition makes the identification conditions testable. The number of skewed shocks is equal to the rank of the coskewness matrix of the reduced-form innovations; likewise, the number of nonmesokurtic shocks is equal to the rank of the cokurtosis matrix of the reduced-form innovations. Guay proposes a bootstrap procedure to implement this test. Keweloh's (2021) identification argument, based on that of Comon (1994), goes through whether skewness conditions, kurtosis conditions, or both are used. An advantage of using skewness conditions alone is that they admit possible coheteroskedasticity in the data, an empirically relevant form of dependence that violates the diagonal coskewness assumptions discussed above.

As originally noted by Bonhomme & Robin (2009), estimation of the higher moments required can be challenging without parametric restrictions due to relatively small sample sizes; this is especially true for inference, which requires up to eighth moments. Guay (2021) uses the identity weighting matrix because of the difficulty of estimating the efficient weighting matrix. For improved finite sample performance, Keweloh (2021) proposes the fast-GMM estimator, which uses a diagonal weighting matrix in which the weight on the covariance moments goes to infinity, so that the shocks are always whitened. The idea is similar to ICA, with the weighting matrix replacing the problem of minimizing dependencies between the shocks with that of maximizing non-Gaussianity, subject to satisfying the covariance moments. In larger models, there is a significant computational advantage to this change, and Keweloh does not find efficiency loss relative to the efficient estimator in simulations. Moreover, he finds that the estimators used by

Lanne & Luoto (2021) and Gouriéroux et al. (2017) do not effectively exploit information found in skewed shocks, instead relying on excess kurtosis, and more generally that information contained in coskewness can be important. Lanne & Luoto (2021) propose two-step, iterative, and continuous updating GMM estimators, as well as a test for overidentifying restrictions and a moment selection procedure based on that of Andrews (1999), although this procedure may struggle in larger SVARs in short samples. In simulations, they prefer the two-step procedure, which in general performs slightly worse than the baseline PML estimator of Gouriéroux et al. (2017) but better than their iterative estimator. Keweloh (2023) then studies the limitations of GMM estimators in short samples. In particular, he notes the challenge of precisely estimating the efficient weighting matrix (which requires up to eighth moments) and also that bias is typically introduced toward parameters implying shock variances below unity. He proposes to impose independence assumptions not only for identification but also in the estimation of the efficient weighting matrix for the former and a continuously updated scaling term for the latter. Simulations show reduced bias and improved coverage. Keweloh (2024) provides a method (discussed in more detail below) to incorporate uncertain economic prior information into moment-based estimators exploiting non-Gaussianity.

In a minimum distance setting, Mesters & Zwiernik (2024) are able to relax the independence assumption in important ways. They show that diagonality of any higher-order cumulant tensor is generally sufficient for identification up to sign and permutation, generalizing away from previous research focused on third and fourth moments (e.g., Bonhomme & Robin 2009, Guay 2021, Keweloh 2021). More importantly, they show that reflectionally invariant restrictions, where the only nonzero cumulant tensor entries are those where each index appears an even number of times, are similarly sufficient for identification. This is the first identification result exploiting non-Gaussianity and fourth moments that can accommodate coheteroskedasticity in the errors, an empirically relevant form of higher-order dependence. However, the shocks must satisfy an additional technical genericity condition. Working in a similar direction, but taking a very different approach, Herwartz & Wang (2024) combine the nonparametric approach of Hafner et al. (2023) with standardization using a kernel estimate of time-varying volatility to identify B in the presence of coheteroskedasticity.

5. WEAK AND UNDERIDENTIFICATION BASED ON HIGHER MOMENTS

It may be tempting to view identification based on higher moments as a free lunch for recovering causal effects in macroeconometrics. No “economic” assumptions are required—at least for local identification—and identification obtains, provided that often uncontroversial properties of the data hold. However, the separation of statistical and economic assumptions is a false dichotomy; assuming that the variances of certain shocks have changed over time in a particular pattern, that all shocks are mutually independent, or that shocks have meaningful excess kurtosis has real economic content. More importantly, the higher moments leveraged by the identification schemes discussed above are all nontrivial to estimate in realistic macroeconomic sample sizes. At best, the schemes exploit the difference in second moments, but more generally they rely on up to the fourth moments of the data, which typically will be very imprecisely estimated in at most 50 years of quarterly data—200 observations.

Researchers are now generally familiar with the weak instruments problem, but similar challenges apply to almost any identification scheme. As described formally by Stock & Wright (2000), when the objective function is relatively flat in the neighborhood of the true parameters, standard inference methods will be invalid, exhibiting large size distortions and poor coverage. This occurs when deviations from homoskedasticity are small, relative to estimation error, so additional values

for the variances offer little information beyond the original $n(n+1)/2$ covariance equations (and deviations of $\hat{B}$ from B are at best weakly rejected by the data in a nontrivial neighborhood of B). Alternatively, when deviations from Gaussianity are small, higher moments offer little information beyond those of a Gaussian distribution—where they are completely redundant with the original covariance restrictions. The possibility that identification based on heteroskedasticity might only be weak—and, thus, that standard inference methods like Wald tests might perform poorly—dates back to at least Magnusson & Mavroeidis (2014), who consider how instability in moments can be used to sharpen identification in general, and propose identification-robust test statistics (Stock & Wright 2000) tailored to such settings. They consider the Rigobon (2003) model as a leading example. More recently, Montiel Olea et al. (2022) argued that weak identification is likely present in many applications of identification based on higher moments.

Lewis (2022) studies the problem of weak identification via heteroskedasticity directly, again in the context of the Rigobon (2003) model. Rigobon & Sack (2004) show that in a bivariate model where only one variance changes that regime-based identification can be rewritten as an IV problem (see Equations 9–11). In that setting, it is unsurprising that weak identification may arise; it does so when the variance that does change does so by only a small amount. In that case, the pretest for weak instruments of Montiel Olea & Pflueger (2013) applies (with bias-based critical values), and the usual robust inference methods for IV estimators can be adopted. Lewis (2022) further characterizes weak identification problems in models identified using variance regimes more generally, which arise when variance changes are close to proportional across regimes. In this general setting, the size-based Andrews (2018) procedure to detect weak identification in GMM problems can be applied. In these models, robust inference is more complicated due to the projection problem: Robust inference can have prohibitively conservative limiting distributions when the econometrician is interested in only a subset of the parameters in B . Lewis (2022) provides conditions under which limiting distributions providing exact size can be derived, which hinge on whether the remaining nuisance parameters can be uniquely determined from the data conditional on the null hypothesis. Empirical evidence suggests that weak identification is present in event studies based on daily financial data, a popular setting in applied practice.

The methods proposed have several limitations. The bias-based pretest is attractive but applies only in a restricted bivariate model. The Andrews (2018) test requires the computation of robust confidence sets and inherits their properties: They can be computationally demanding to construct in larger models and prohibitively conservative when the conditions provided for improved limiting distributions do not hold. Finally, those conditions place strong limits on the extent of weak identification, which may be hard to justify for $n > 2$.

Lütkepohl et al. (2021) provide a test of the identification condition for the Rigobon (2003) identification scheme using two regimes. They test whether equality can be rejected for each subset of eigenvalues in Equation 7. This is not a test for weak identification, but rather a test of non- or underidentification. A limitation of this test is that it presents a potentially substantial multiple testing problem when $n > 2$: To reject nonidentification, as many tests as there are subsets of adjacent integers in $1, \dots, n$ must be conducted. However, it remains an attractive option in applied practice, given the computational and performance issues of the Andrews (2018) procedure for larger models. Bacchiocchi & Kitagawa (2020) develop a Bayesian approach that defaults to robust Bayesian inference (as in Giacomini & Kitagawa 2021) and treat the model as set-identified whenever weak identification is a concern [e.g., when the Lütkepohl et al. (2021) test indicates that the eigenvalues of Equation 7 are not distinct].

Similar tests for nonidentification exist for other schemes. Lewis (2021) derives testable implications of his distribution-free identification conditions and proposes a suitable Cragg–Donald statistic. Lanne & Saikkonen (2007) propose two Lagrange Multiplier-type tests for the order

of the GARCH process driving the SVAR innovations, which can help determine whether the required $n - 1$ dimensions are present or whether identification is only partial. Lütkepohl & Milunovich (2016) propose an additional test and study all three in extensive simulations. The size and power of the tests and their ranking varies with data-generating processes and sample sizes, and they tend to be conservative in determining the dimension of heteroskedasticity. However, Lanne & Saikkonen (2007) and Lütkepohl & Milunovich (2016) consider only the case where there are r heteroskedastic shocks and $n - r$ homoskedastic shocks, rather than models where all shocks may be heteroskedastic, but with a factor structure in the volatilities. Lütkepohl et al. (2020) consider several tests for the rank of heteroskedasticity in order to study partial identification.

Turning to non-Gaussianity, serious effort has recently been applied to robust inference. Lee & Mesters (2021) and Hoesch et al. (2024) provide a robust inference approach based on a semiparametric score statistic (with the nonparametric part corresponding to the shock distributions). Lee & Mesters (2021) consider identification via non-Gaussianity from observed innovations, while Hoesch et al. (2024) extend the results to SVARs and IRFs and show how the score can be used to construct an efficient estimator. Inference on IRFs ultimately requires a Bonferroni step and projection methods. The confidence intervals constructed by Drautzburg & Wright (2021), based on sign restrictions and an independence assumption, are also robust to weak identification (with respect to the variation from non-Gaussianity).

Recently, Amengual et al. (2022a,b) provided tests of the assumptions for identification based on non-Gaussianity, based on the estimated shocks. The authors propose moment-based tests for whether one or more shocks has a Gaussian distribution and for dependence in the shocks (Amengual et al. 2022a) and present a test for dependence among the shocks that compares the estimated joint cumulative distribution function of the shocks with the estimated marginal cumulative distribution functions (Amengual et al. 2022b). Both tests are based on the mixture of normals PML estimator of Fiorentini & Sentana (2023), for which influence functions allow adjustment for estimation error in the shocks. Additional research on testing the independence assumption alone can be found in the ICA literature (e.g., Matteson & Tsay 2017, Davis & Ng 2023).

Notably, proposals to pretest for Gaussianity are relatively rare, given that non-Gaussianity is a necessary condition for any identification approach based on higher moments. A finding of Gaussian shocks can preclude identification entirely or limit the scope to partial identification. As discussed above, Amengual et al. (2022a) and Maxand (2020) recently developed tests for Gaussianity of shocks in the context of identification based on non-Gaussianity in moment-based and maximum likelihood settings, respectively. However, consideration of such tests in the literature is limited, and they are rarely adopted in practice. In contrast, there is a developed literature on testing for Gaussianity outside of identification based on higher moments, which could be exploited. For instance, Bai & Ng (2005) propose tests for skewness, kurtosis, and non-Gaussianity more generally tailored to time-series settings; they find testing for skewness in particular to be effective.

Finally, while empirical researchers might find guidance helpful on how large a sample is required to obtain reliable estimates based on higher moments, there is currently no sound basis for such guidance. Rather than clear recommendations, the overwhelming message of simulation studies (e.g., Lewis 2021, Herwartz et al. 2022, Keweloh 2023) is that, for a given sample size, performance can vary starkly with the degree of parameterization imposed in addition to the restrictions required for identification (and, of course, whether those additional assumptions are consistent with the data). Of course, the effective sample size will always depend on the strength of the identifying variation and any dependence in the data as well. These studies further illustrate that, even absent additional parametric structure, results may vary dramatically with the precise

estimator used for a given sample size—for instance, one based on variance regimes as opposed to one based on unconditional higher moments.

6. OPEN ISSUES

In this section, I review several open issues and areas of ongoing research related to identification based on higher moments. In particular, I address the possibility of time variation in B , the trade-offs of distribution-free versus likelihood-based approaches, the combination of different sources of identifying information, and the shock labeling problem that arises due to local identification up to only column permutations (and possibly signs of the shocks).

6.1. Time-Varying B

A possible tension in standard models identified via heteroskedasticity is that the structural parameters in B —the causal effects—are required to be constant over time, while the variances are allowed to change. Indeed, these two sources of variation are sometimes included together in models where identification comes from other sources (e.g., Cogley & Sargent 2005, Primiceri 2005). On the other hand, some dynamic stochastic general equilibrium models include time-varying volatility but not changes in the deep structural parameters that dictate B in a VAR representation. In any case, arguments for identification based on higher moments hinge on the constancy of B . However, it is possible to accommodate time-varying reduced-form parameters, $A(L)$, and the failure to do so when demanded by the data can be a source of spurious heteroskedasticity in SVAR innovations (Sims 2002). There have been several recent attempts to combine identification via heteroskedasticity with particular forms of time variation in B . Typically, these approaches make use of additional regimes that would be overidentifying with constant coefficients in order to identify changes in a restricted number of coefficients. Note, however, that adding further regimes cannot identify time-varying B and volatilities without such restrictions: Each additional regime adds n new volatility parameters and n^2 new coefficients, but only $(n^2 + n)/2$ new equations. Often, the identification results in this literature provide local identification conditions under which the Jacobian of the reduced-form covariance matrices with respect to the unrestricted parameters is full rank.

Bacchiocchi & Fanelli (2015) and Bacchiocchi et al. (2018) propose a framework that nests that of Rigobon (2003). They consider two regimes with the covariance matrix of the shocks constant but with two matrices of coefficients, B and $\tilde{B} \equiv B + W$, modeled as functions of a vector of $n^2 + n$ unknown parameters. Note that changes in shock variances can be subsumed into $\tilde{B} = B + W$, where each column is rescaled only by the new volatility additively instead of multiplicatively. They derive the rank condition for identification; unsurprisingly, changes in relative effects of the shocks can be identified only if some of the variances are restricted to be unchanged (via the matrix W). These are the same equations that were just-identifying with B fixed and all volatilities allowed to change. Angelini et al. (2019) go a step further. They likewise do not explicitly model heteroskedasticity but rather exploit it to identify three different regimes of impact coefficients in a study of economic uncertainty. In particular, they keep the shock variances fixed at unity and model the impact coefficients as B , $B + W_1$, and $B + W_1 + W_2$. This structure allows for changes in the volatilities via W_1 and W_2 , as well as changes in the relative effects of the shocks. Identification is achieved by imposing a suitable number of zero restrictions in W_1 and W_2 . Recently, Bacchiocchi & Kitagawa (2023) considered more general SVARs with breaks permitted in all reduced-form and structural parameters, where the econometrician may exploit stability restrictions, inequalities across regimes, standard zero and sign restrictions, and constraints on forecast error variance decompositions. They provide conditions for point and set identification. Brenna et al. (2023) take a

different perspective. Studying macrofinancial linkages, they note that there are too few uncontroversial restrictions that can be imposed on time variation in B to offer point identification. Instead, for a regime $i > 1$ they let $B_i = B + W_i$, construct the full set of combinations of zero restrictions that could be sufficient to identify the model, and recover the identified set of parameters across that collection of sets of possible identifying restrictions.

Dungey et al. (2015) marry the smooth transition and GARCH-based identification approaches. They model volatilities as evolving according to a GARCH process and the time-varying B_t as a convex combination of $N + 1$ unit diagonal matrices:

$$B_t = (1 - S_N) [\dots [(1 - S_2) [(1 - S_1) B_0 + S_1 B_1] + S_2 B_2] \dots] + S_N B_N. \quad 23.$$

Here, $S_j = [1 + e^{-\gamma(x_t - d_j)}]^{-1}$, where d_j is the center of the transition between regimes $j - 1$ and j and $x_t = t/T$ expresses time t as a fraction of the sample. The identification argument contends that when the speed of transition, γ , is large enough that within each regime j , B_t is essentially B_j , then identification results like that of Milunovich & Yang (2013) can be applied regime by regime. However, the true parameters appear to be identified only in the limit as $\gamma \rightarrow \infty$. Camehl & Woźniak (2024) propose to combine a Markov switching process for economically plausible exclusion restriction patterns for monetary policy shocks with stochastic volatility, allowing for time variation in B ; however, formal joint identification results for the regimes and shock volatilities appear challenging.

6.2. Choosing the Right Functional Form

When choosing from among identification approaches based on heteroskedasticity, the econometrician is usually required to take a stance on the form of heteroskedasticity, at least to some extent. While now behind the frontier range of options for the volatility process, Lütkepohl & Netšunajev (2015) provide an early survey of possible models, highlighting their relative strengths and weaknesses. All identification schemes (except for that in Lewis 2021) rely on estimating the paths of the reduced-form variances through time, at least for feasible implementations. While the use of a regime approach can be thought of as unrestrictive—since variance regimes can be estimated through time regardless of the true variance process—they will offer sharp identifying information only if there are distinct differences in structural variances across regimes. If the true variance process is poorly approximated by a regime structure, it is unlikely that identification will be strong and, at the very least, valuable identifying information will be left on the table. However, even though distribution-free estimators are available under fairly weak assumptions, it may be preferable to specify a functional form for the variance process in finite samples for efficiency in estimating higher moments.

Therefore, how should the researcher go about choosing the variance process to fit to the data? This remains an open question, with three possible answers. The first option is to use application-specific knowledge to choose an appropriate model. The second is to conduct statistical tests to learn the correct functional form, as proposed by Lütkepohl & Schlaak (2018), who provide various information criteria for determining the correct model of heteroskedasticity. They then evaluate their performance in a simulation study. They consider exogenously determined regimes, Markov switching, smooth transition, and GARCH DGPs. They find that the information criteria struggle to differentiate between different models for the structural variances, in terms of making the correct binary determination. They tend to favor the exogenous regime model and, in particular, struggle to detect GARCH data. However, adopting the criteria can still help reduce the mean squared error of impulse response estimates. Ultimately, they conclude that these tests can be helpful in ascertaining whether adequate heteroskedasticity is present in the data for identification, but they are not yet well-suited for discriminating among different volatility models.

The third option is to choose an estimator to be as robust as possible. Lewis (2021) and Bertsche & Braun (2022) conduct parallel simulation studies comparing the performance of various different estimators (in terms of MSE and other criteria) when correctly and incorrectly specified, pointing to the possibility of choosing a demonstrably robust functional form. Specifically, they consider estimators based on AR(1) log stochastic volatility, GARCH, Markov switching, and regime-based models for the structural variances, as well as a distribution-free estimator based on GMM. Lewis (2021) adds further estimators, including two exploiting non-Gaussianity. Both studies consider DGPs corresponding to each of these models, and both find that the AR(1) log stochastic volatility model is remarkably robust to misspecification. It performs very well when correctly specified, and even when badly misspecified it often performs nearly as well as correctly specified estimators. This is not true for any of the other estimators, whose performance is generally very heterogeneous. Both studies recommend the stochastic volatility model for use in practice on the basis of its robustness to misspecification. However, beyond the simple flexibility of the DGP (ex post, paths for the latent volatilities can approximate those coming from any of the other variance processes, and even the fat tails of homoskedastic non-Gaussian DGPs), no theoretical explanation for this performance or justification for this guidance is available to date. Separately, both studies find that the simple GMM implementation of the Lewis (2021) scheme performs poorly in realistic sample sizes as a result of the difficulty in estimating the higher moments required.

For the case of non-Gaussianity, similar questions arise, and unfortunately there is little clear guidance for empirical researchers. There are maximum likelihood and PML approaches, typically based on variations on Student t distributions, ICA-type nonparametric estimators, and newer moment-based approaches. For the last, an important choice is which coskewness and/or cokurtosis restrictions to impose, or which contrast function to use for ICA. Moneta & Pallante (2022) compare a variety of estimators in a simulation study. They include FastICA, the PML estimator of Gouriéroux et al. (2017), and two other ICA approaches based on Givens matrices; unfortunately, they do not include recent moment-based estimators. Overall, they find that FastICA has an edge in terms of bias, efficiency, size distortions, and coverage, with PML slightly behind. Herwartz et al. (2022) actually compare both non-Gaussianity-based schemes and heteroskedasticity-based identification in a somewhat wider simulation study based on a range of DGPs. They consider the maximum likelihood estimator of Lanne et al. (2017) and nonparametric ICA approaches as well as regime- and GARCH-based methods but, again, none of the recent moment-based approaches. While all estimators generally perform well when well-specified, a nonparametric “dCov” ICA approach from Matteson & Tsay (2017) alone performs consistently well across DGPs. As discussed in Section 4.4, several studies have noted the challenges involved in estimating the required moments in realistically small sample sizes, similar to those noted for Lewis’s (2021) moment-based approach. These open questions make an exhaustive comparison with alternative estimators of great interest and provides fruitful ground for refinements to estimation of the moments, as considered by, for example, Keweloh (2023).

6.3. Combining Multiple Sources of Identifying Information

An exciting avenue for ongoing research focuses on combining identification based on higher moments with other identification schemes. This approach can serve three purposes. First, additional overidentifying assumptions—particularly economic ones—can be tested. Second, the combination of statistical identifying information with other types of identifying information, like an external instrument, zero restrictions, or sign restrictions, can serve to sharpen identification when the heteroskedasticity or non-Gaussianity may not be pronounced enough to provide strong identification in finite samples, or may offer only partial identification. Finally, it can help resolve the labeling indeterminacy (for discussion, see Section 6.4).

The first purpose is one of the key advantages of statistical identification in the first place. For example, B can be estimated on the basis of higher moments alone, and then a simple joint Wald test for all of the entries above the diagonal can determine whether recursive (short-run) identification assumptions are rejected by the data. Alternatively, the same can be done using likelihood ratio tests and other approaches. Myriad examples of this approach to testing economic identification assumptions exist, including the empirical applications of virtually all of the references in this review (e.g., Normandin & Phaneuf 2004, Lanne & Lütkepohl 2008, Herwartz & Lütkepohl 2014, Lütkepohl & Woźniak 2020, Lewis 2021, Bertsche & Braun 2022).

Combining higher moments with additional identifying assumptions to achieve sharper identification—for example, when one fears that the higher moments may only be weakly identifying—is still an emerging literature. Carrierio et al. (2023) provide a wide-ranging treatment in a Bayesian framework. They propose algorithms to estimate SVARs that combine heteroskedasticity with sign and narrative restrictions as well as external instruments. They argue that heteroskedasticity, since it is potentially point-identifying, can substantially reduce the identified sets resulting from sign or narrative restrictions. In contrast, those restrictions can resolve the labeling problem associated with statistical identification. A key challenge of such strategies is understanding how conflicting identifying information will interact. Indeed, the identified set may be empty if two sources of identifying information are at odds—for instance, if no shock exists, according to the moments arising from heteroskedasticity, that satisfies stipulated sign restrictions. In that case, determining which set of assumptions is incorrect presents a further challenge. Lütkepohl & Schlaak (2022) consider a slightly different problem. They take the presence of heteroskedasticity (in regimes, say) as potential evidence of time-varying impact coefficients. They combine heteroskedasticity and an external instrument to test whether the column of B identified by the instrument changes across variance regimes. Bacchiocchi et al. (2024) derive identified sets resulting from partially identifying heteroskedasticity and zero or sign restrictions and provide methods to compute them as well as methods for robust Bayesian inference. Bacchiocchi & Kitagawa (2023) provide a comprehensive treatment of identification in SVARs with breaks, which nests models identified with heteroskedasticity regimes but also allows for breaks in B and the reduced-form parameters. They allow for additional information in the form of inequalities on various SVAR objects and stability restrictions and provide Bayesian and robust Bayesian algorithms and methods for valid inference.

Drautzburg & Wright (2021) combine sign restrictions with non-Gaussianity. They propose an identified set that is the intersection of that arising from sign restrictions and the set of models for which independence of the shocks cannot be rejected. Andrade et al. (2024) take a different approach. They impose not independence but rather only sign restrictions on third or fourth moments, in addition to economic sign restrictions, to sharpen the identified set. Keweloh et al. (2023a) combine a block recursive structure with non-Gaussian shocks to propose an estimator that has advantages over purely statistical information in terms of performance, shock labeling, and weaker independence requirements. Braun (2021) develops a Bayesian framework combining priors involving sign restrictions and non-Gaussianity. He models each non-Gaussian shock as a univariate Dirichlet process mixture model and shows that, combined with weak priors on important coefficients, the non-Gaussianity sharpens inference to deliver results similar to those under much stronger restrictions in an application to oil markets. Keweloh et al. (2023b) propose a Bayesian framework that allows for potentially endogenous proxy variables and show that that such endogeneity helps to reconcile the range of empirical estimates for fiscal multipliers. Keweloh (2024) combines non-Gaussianity with potentially invalid short-run restrictions, proposing an estimator with data-dependent shrinkage toward those restrictions. Herwartz & Wang (2023) develop a point estimator that minimizes the dependence of the implied structural shocks

subject to sign restrictions. Crucil et al. (2023) exploit external instruments with non-Gaussianity to sharpen identification and partially resolve the shock labeling problem; the relative importance of the features depends on the strength of the instruments.

6.4. The Shock Labeling Problem

One of the main challenges of purely statistical identification is the shock labeling problem— B is only ever identified up to column order, and possibly also column sign, by statistical information alone.⁵ The solution, in general, is to use economic information to label the shocks (for an early detailed discussion, see Herwartz & Lütkepohl 2014). While this may seem self-defeating, having turned to statistical identification to avoid economic assumptions, the versions of economic assumptions required for labeling are generally much less restrictive than those required for point identification on their own. For example, rather than imposing a lower-triangular structure on B , the researcher can obtain the permutation of the identified columns of B that is “closest” to that structure, under the Frobenius norm, say. The shock that is best-correlated with a proxy variable can be chosen without assuming that the proxy is exogenous. The economic assumptions do not determine the possible values of identified parameters—they are recovered on statistical information alone; rather, the economic information helps in choosing among the observationally equivalent statistically identified models, which is desirable if the economic restrictions are thought to be approximately but perhaps not literally true. Another alternative is to obtain a set of model-based causal effects (values for B , or IRFs at longer horizons) and choose the shock labeling that most closely matches the model predictions (e.g., Brunnermeier et al. 2021). Augmenting the statistical identification assumptions with some economic assumption that selects a unique column order (or even selects a single column of interest) is enough to point-identify the causal effects of interest. Lewis (2021) and Kilian & Lütkepohl (2017, chapter 14) discuss the labeling problem in further detail. Of course, sometimes the economic assumptions supplied may not be adequate to achieve point identification among permutations, and identification remains local (for a discussion, see Bacchiocchi & Kitagawa 2020).

There are also noneconomic options for pinning down a unique value for B , beyond column order. For example, Lanne et al. (2017) propose to globally identify an SVAR based on non-Gaussianity using a series of transformations. Given any one of the $n!$ observationally equivalent identified matrices, they first normalize each column to have unit Euclidean norm. Next, they choose the unique column permutation for which the entries to the right of each diagonal element are smaller in absolute value than that diagonal element. Finally, they impose the unit diagonal normalization. That can be thought of as an approximation to a lower-triangular matrix, since one chooses the permutation chosen for which all of the entries above the diagonal are smaller than the corresponding diagonal entries. Another popular choice in the ICA literature, and in recent papers using non-Gaussianity, is the Pham & Garat (1997) scheme, which chooses the permutation that maximizes the product of the diagonal entries of B , after restricting them to be positive and making a unit variance assumption. While augmenting statistical assumptions with such column ordering rules achieves global identification, it does not directly assist in rendering identified shocks (or their effects) economically interpretable.

An additional concern is the impact of the labeling problem on inference. While the asymptotic distribution of a given column permutation of B may be known (after fixing some statistical rule, like those mentioned above, for estimation), what is the distribution after subsequent permutations and rescaling for economic labeling? Lewis (2021) applies results from the model

⁵The shock labeling problem results in the challenging label switching problem for Bayesian inference (for a discussion, see, e.g., Bacchiocchi & Kitagawa 2020).

selection literature to show that, provided a consistent labeling criterion is used, the error introduced by the labeling procedure is asymptotically negligible, and the asymptotic distribution of the permutation of interest is valid for the labeled estimator without modification. A consistent labeling criterion is a rule for choosing the column order of B that will choose the economically correct order with probability approaching one as $T \rightarrow \infty$. This result meaningfully simplifies inference on the parameters of interest in practice.

7. CONCLUSION

To conclude, I review several areas fruitful for ongoing research.

7.1. Robust Inference

As discussed above, higher moments often likely provide only weakly identifying information in realistic sample sizes. Although this fact has been noted in the theoretical literature for some time, it has only recently started to influence applied practice. In general, weakly identified SVARs require projection inference when the researcher is interested in a subset of the parameters, with available critical values potentially prohibitively conservative, as observed by Lewis (2022) for heteroskedasticity and Lee et al. (2022) for non-Gaussianity. Lewis (2022) provides a solution for a particular class of models in the form of considerably smaller critical values, but for most applications no such results exist. Robust inference methods that exhibit attractive performance and are computationally convenient remain elusive and, as the literature comes to grips with the prevalence of weak identification, will increasingly be in demand.

7.2. Testing Identification Conditions

Broadening the range of settings in which the identification conditions are testable should be a priority. Since the identification conditions are in terms of the structural parameters, it is challenging to test them without assuming identification in the first place. Doing so generally requires the derivation of testable implications in terms of only reduced-form quantities. Several methods are discussed in Section 5. Testing approaches, especially those testing for weak identification as opposed to nonidentification, that perform well and are computationally tractable for arbitrarily large models remain a target for all of the identification schemes considered above. In another vein, there is ground for further research to develop inference approaches under partial identification, which could be detected by such tests.

7.3. Combining Identifying Information

As discussed in Section 6.3, several recent papers have combined statistical identification with other forms of identification, aiming to combat the weakness of the identifying variation coming from higher moments, provide point identification when economic restrictions are only set-identifying, or test economic restrictions. This is a conceptually appealing strategy, since it addresses head-on the fact that many sources of identifying information fail to provide sharp identification, or may offer only partial identification. It has the potential to exploit the best of both worlds in terms of statistical and economic identifying information. The Bayesian literature may be somewhat further advanced in this respect because it is particularly appealing to combine higher moments with set identification approaches, which are often implemented in a Bayesian framework. However, there is plenty of scope to pursue such developments in the frequentist paradigm as well (e.g., Drautzburg & Wright 2021, Keweloh 2024). In particular, a comprehensive framework exploiting information from higher moments, external instruments, zero restrictions,

and possibly sign restrictions, including methods to test the inevitably complicated identification condition(s), should be the goal. Ongoing research by Carriero et al. (2021) and Bacchiocchi & Kitagawa (2023) appears to come the closest to this goal, in a Bayesian context. In addition to theoretical research, this area represents a rich vein for empirical studies, since the combination of all credible identifying information available will likely pay dividends.

7.4. Functional Forms

Section 6.2 outlines the challenges of choosing the best functional form—and estimation approach more generally—for a particular application. For both identification based on heteroskedasticity and identification based on non-Gaussianity, there are meaningful decisions for the researcher to make in terms of whether to choose a nonparametric estimator or specify some particular functional form for the volatility process or distribution for the shocks. Parametric estimators can be particularly appealing due to the difficulty in precisely estimating higher moments in realistic macroeconomic sample sizes. As discussed above, recent attempts to detect the best-fitting form of heteroskedasticity based on information criteria have proven unsatisfactory (Lütkepohl & Schlaak 2018). Lütkepohl & Schlaak's (2018) study also predates the availability of identification results for a much richer range of functional forms, like the AR(1) log stochastic volatility model favored by Bertsche & Braun (2022) and Lewis (2021). Thus, there is scope both to extend Lütkepohl & Schlaak's (2018) analysis and to investigate alternative testing methodologies that may offer better performance.

To my knowledge, there has been no systematic and comprehensive comparison of the quickly multiplying options for implementation of identification via non-Gaussianity (e.g., one including moment-based approaches). The speed at which new implementations of identification based on non-Gaussianity are developed makes such comparisons difficult; current options described in Section 4 include various parametric models, moment-based approaches where the researcher must choose which higher moments to use and which restrictions implied by independence to impose, and a plethora of nonparametric estimators from the ICA literature. Moneta & Pallante (2022) and Herwartz et al. (2022) present the only simulation comparisons of which I am aware. There is ample scope for more comprehensive simulation studies comparing the leading alternatives, as well as “pretests” to select suitable moments containing relevant identifying variation.

Separately, there is also scope for creative innovation under each of these identification approaches. For example, Lewis (2021) justifies identification under a very wide class of persistent volatility models. This frees the researcher to exploit essentially whatever persistent volatility process she thinks best suits the data, providing substantial ground for exploration. In the case of non-Gaussianity, while several early parametric implementations started from obvious fat-tailed distributions, the identification requirements likewise admit a continuum of possible distributions, and there is scope to propose new alternatives here as well. For example, Jarociński (2024) proposes a novel likelihood that allows for dependence in the tails of the shocks, motivated by empirical features of US monetary policy. Moreover, Mesters & Zwiernik (2024) provide exciting new results that relax the independence assumption and open the door for researchers to work on a much richer and more realistic range of distributions (including, e.g., those with coheteroskedasticity). Alternatively, it may be worth incorporating shrinkage techniques to estimate higher moments, from which to identify the SVAR parameters.

7.5. Presenting Statistical Identification to Applied Researchers

While statistical identification often exploits relatively uncontroversial properties of the data, the mechanisms of identification can be much less clear. Many applied researchers want to know what economic features identify the parameters. Statistical properties of the data do have economic

meaning. Unpacking statistical identification in economic terms can be challenging and is an area where econometricians must do better. The original paper by Rigobon (2003) is exemplary in this respect: It represents identification based on variance regimes graphically and makes clear how identification obtains as the variances of two shocks change across them. Lewis (2021) attempts to reframe rather abstract conditions on the persistence and copersistence of shock variances in terms of predictions from structural models and conceptual properties of fiscal shocks, as well as interpreting periods of high and low volatility and particularly sizable shocks through the lens of historical events. While it is natural to use information from the historical record to define variance regimes, there is no reason such narrative information cannot also be used to interpret the variation underlying any other approach using higher moments.

Similar arguments can be made with respect to non-Gaussianity. While somewhat more challenging, theory can also help motivate the required independence or other restrictions on higher moments necessary to achieve identification. Guay (2021) provides a compelling graphical illustration of the identification argument, mirroring that of Rigobon (2003) for heteroskedasticity. Braun (2021) argues similarly, illustrating the power of non-Gaussianity for identifying supply and demand shocks in oil markets. Greater effort along these lines to build intuition for what drives identification, aside from a collection of equations and associated rank conditions, will increase applied uptake and the appeal of future research in these areas.

Statistical identification exploiting higher moments offers tremendous potential as an alternative or complementary source of identifying information, alongside economic restrictions. However, identification does not come for free, and it is important to understand the limitations of these approaches. Over the past two decades, the literature has developed a rich range of identification schemes, and it is now beginning to assess and address the associated challenges in earnest. Research in the coming years should aim to provide empirical researchers with the tools they need to understand and judiciously select an appropriate combination of identifying assumptions, formally evaluate the validity of those assumptions, rigorously choose among available implementations, conduct reliable inference on their estimates, and present their findings in an intuitive and compelling manner.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I am grateful to Emanuele Bacchiocchi and Sascha Keweloh for helpful comments and to Brian Amorim Cabaço and Ziyu Peter Jiang for research assistance.

LITERATURE CITED

- Amengual D, Fiorentini G, Sentana E. 2022a. Moment tests of independent components. *SERIEs J. Span. Econ. Assoc.* 13(1):429–74
- Amengual D, Fiorentini G, Sentana E. 2022b. *Specification tests for non-Gaussian structural vector autoregressions*. Work. Pap. 2022-2212, Cent. Monet. Financ. Stud., Madrid
- Andrade P, Ferroni F, Melosi L. 2024. *Higher-order moment inequality restrictions for SVARs*. Tech. Rep. 2023-28, Fed. Reserve Bank, Chicago
- Andrews DWK. 1999. Consistent moment selection procedures for generalized method of moments estimation. *Econometrica* 67(3):543–63
- Andrews I. 2018. Valid two-step identification-robust confidence sets for GMM. *Rev. Econ. Stat.* 100(2):337–48

- Angelini G, Bacchiocchi E, Caggiano G, Fanelli L. 2019. Uncertainty across volatility regimes. *J. Appl. Econom.* 34(3):437–55
- Anttonen J, Lanne M, Luoto J. 2023. *Bayesian inference on fully and partially identified structural vector autoregressions*. Work. Pap., Univ. Helsinki, Helsinki, Finl.
- Bacchiocchi E, Bastianin A, Kitagawa T, Mirto E. 2024. *Partially identified heteroskedastic SVARs*. Work. Pap. 15.2024, Fond. Eni Enrico Mattei, Milan, Italy
- Bacchiocchi E, Castelnovo E, Fanelli L. 2018. Gimme a break! Identification and estimation of the macroeconomic effects of monetary policy shocks in the U.S. *Macroecon. Dyn.* 22(6):1613–51
- Bacchiocchi E, Fanelli L. 2015. Identification in structural vector autoregressive models with structural changes, with an application to US monetary policy. *Oxford Bull. Econ. Stat.* 77(6):761–79
- Bacchiocchi E, Kitagawa T. 2020. *Locally- but not globally-identified SVARs*. Work. Pap. 40/20, Cent. Microdata Methods Pract., London
- Bacchiocchi E, Kitagawa T. 2023. SVARs with breaks: identification and inference. arXiv:2405.04973 [econ.EM]
- Bai J, Ng S. 2005. Tests for skewness, kurtosis, and normality for time series data. *J. Bus. Econ. Stat.* 23(1):49–60
- Bernstein S. 1941. On a property which characterizes a Gaussian distribution. *Proc. Leningrad Polytech. Inst.* 217(3):21–22
- Bertsche D, Braun R. 2022. Identification of structural vector autoregressions by stochastic volatility. *J. Bus. Econ. Stat.* 40(1):328–41
- Blanchard OJ, Quah D. 1989. The dynamic effects of aggregate demand and supply disturbances. *Am. Econ. Rev.* 79(4):655–73
- Blundell R, Pistaferri L, Preston I. 2008. Consumption inequality and partial insurance. *Am. Econ. Rev.* 98(5):1887–921
- Bonhomme S, Robin JM. 2009. Consistent noisy independent component analysis. *J. Econom.* 149(1):12–25
- Bouakez H, Normandin M. 2010. Fluctuations in the foreign exchange market: How important are monetary policy shocks? *J. Int. Econ.* 81(1):139–53
- Braun R. 2021. *The importance of supply and demand for oil prices: evidence from non-Gaussianity*. Work. Pap. 957, Bank Engl., London
- Brenna F, de Graeve F, Wouters R. 2023. *Macro-financial feedbacks through time*. Work. Pap., KU Leuven/Natl. Bank Belg.
- Brüggemann R, Jentsch C, Trenkler C. 2016. Inference in VARs with conditional heteroskedasticity of unknown form. *J. Econom.* 191(1):69–85
- Brunnermeier M, Palia D, Sastry KA, Sims CA. 2021. Feedbacks: financial markets and economic activity. *Am. Econ. Rev.* 111(6):1845–79
- Camehl A, Woźniak T. 2024. Time-varying identification of monetary policy shocks. arXiv:2502.19659 [econ.EM]
- Cardoso J, Souloumiac A. 1993. Blind beamforming for non-Gaussian signals. *IEEE Proc. F* 140(6):362–70
- Carriero A, Clark T, Marcellino M. 2021. Using time-varying volatility for identification in vector autoregressions: an application to endogenous uncertainty. *J. Econom.* 225(1):47–73
- Carriero A, Marcellino M, Tornese T. 2023. *Blended identification in structural VARs*. Work. Pap. 23200, Cent. Appl. Res. Int. Mark. Finance Regul., Univ. Bocconi, Milan, Italy
- Chen A, Bickel PJ. 2006. Efficient independent component analysis. *Ann. Stat.* 34(6):2825–55
- Cogley T, Sargent TJ. 2005. Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Rev. Econ. Dyn.* 8(2):262–302
- Comon P. 1994. Independent component analysis, a new concept? *Signal Proc.* 36(3):287–314
- Crucil R, Hambuckers J, Simone M. 2023. *Do monetary policy shocks affect financial uncertainty? A non-Gaussian proxy SVAR approach*. Work. Pap., Univ. Liège/Eur. Univ. Viadrina
- Darmois G. 1953. Analyse générale des liaisons stochastiques: étude particulière de l'analyse factorielle linéaire. *Rev. Inst. Int. Stat./Rev. Int. Stat. Inst.* 21(1/2):2–8
- Davis R, Ng S. 2023. Time series estimation of the dynamic effects of disaster-type shocks. *J. Econom.* 235(1):180–201
- Drautzburg T, Wright JH. 2021. *Refining set-identification in VARs through independence*. Work. Pap. 64575, Johns Hopkins Univ., Baltimore, MD

- Dungey M, Milunovich G, Thorp S, Yang M. 2015. Endogenous crisis dating and contagion using smooth transition structural GARCH. *J. Bank. Finance* 58:71–79
- Fernández-Villaverde J, Rubio-Ramírez JF, Sargent TJ, Watson MW. 2007. ABCs (and Ds) of understanding VARs. *Am. Econ. Rev.* 97(3):1021–26
- Fiorentini G, Sentana E. 2023. Discrete mixtures of normals pseudo maximum likelihood estimators of structural vector autoregressions. *J. Econom.* 235(2):643–65
- Geary RC. 1941. Inherent relations between random variables. *Proc. R. Irish Acad. A* 47:63–76
- Giacomini R, Kitagawa T. 2021. Robust Bayesian inference for set-identified models. *Econometrica* 89(4):1519–56
- Gospodinov N, Komunjer I, Ng S. 2017. Simulated minimum distance estimation of dynamic models with errors-in-variables. *J. Econom.* 200(2):181–93
- Gospodinov N, Ng S. 2015. Minimum distance estimation of possibly noninvertible moving average models. *J. Bus. Econ. Stat.* 33(3):403–17
- Gouriéroux C, Monfort A, Renne JP. 2017. Statistical inference for independent component analysis: application to structural VAR models. *J. Econom.* 196(1):111–26
- Gouriéroux C, Monfort A, Renne JP. 2019. Identification and estimation in non-fundamental structural VARMA models. *Rev. Econ. Stud.* 87(4):1915–53
- Guay A. 2021. Identification of structural vector autoregressions through higher unconditional moments. *J. Econom.* 225(1):27–46
- Hafner CM, Herwartz H, Wang S. 2023. *Causal inference with (partially) independent shocks and structural signals on the global crude oil market*. Discuss. Pap. 2023004, KU Leuven, Leuven, Belg.
- Herwartz H, Lange A, Maxand S. 2022. Data-driven identification in SVARs—when and how can statistical characteristics be used to unravel causal relationships? *Econ. Inq.* 60(2):668–93
- Herwartz H, Lütkepohl H. 2014. Structural vector autoregressions with Markov switching: combining conventional with statistical identification of shocks. *J. Econom.* 183(1):104–16
- Herwartz H, Wang S. 2023. Point estimation in sign-restricted SVARs based on independence criteria with an application to rational bubbles. *J. Econ. Dyn. Control* 151:104630
- Herwartz H, Wang S. 2024. *Consistent statistical identification of SVARs under (co-)heteroskedasticity of unknown form*. Work. Pap., Univ. Göttingen, Göttingen, Ger.
- Hoesch L, Lee A, Mesters G. 2024. Locally robust inference for non-Gaussian SVAR models. *Quant. Econ.* 15(2):523–70
- Hyvärinen A. 1997. New approximations of differential entropy for independent component analysis and projection pursuit. *Adv. Neural Inf. Proc. Syst.* 10:273–79
- Hyvärinen A. 1999. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Netw.* 10(3):626–34
- Hyvärinen A. 2013. Independent component analysis: recent advances. *Philos. Trans. R. Soc. A* 371(1984):20110534
- Hyvärinen A, Karhunen J, Oja E. 2001. ICA by maximum likelihood estimation. In *Independent Component Analysis*, pp. 203–19. New York: Wiley
- Hyvärinen A, Zhang K, Shimizu S, Hoyer PO. 2010. Estimation of a structural vector autoregression model using non-Gaussianity. *J. Mach. Learn. Res.* 11(56):1709–31
- Jarociński M. 2024. Estimating the Fed's unconventional policy shocks. *J. Monet. Econ.* 144:103548
- Kac M. 1939. On a characterization of the normal distribution. *Am. J. Math.* 61(3):726–28
- Karamysheva M, Skrobotov A. 2022. Do we reject restrictions identifying fiscal shocks? Identification based on non-Gaussian innovations. *J. Econ. Dyn. Control* 138:104358
- Keweloh SA. 2021. A generalized method of moments estimator for structural vector autoregressions based on higher moments. *J. Bus. Econ. Stat.* 39(3):772–82
- Keweloh SA. 2023. Structural vector autoregressions and higher moments: challenges and solutions in small samples. arXiv:2310.08173 [econ.EM]
- Keweloh SA. 2024. Uncertain short-run restrictions and statistically identified structural vector autoregressions. arXiv:2303.13281 [econ.EM]
- Keweloh SA, Hetzenecker S, Seepa A. 2023a. Monetary policy and information shocks in a block-recursive SVAR. *J. Int. Money Finance* 137:102892

- Keweloh SA, Klein M, Prüser J. 2023b. Estimating fiscal multipliers by combining statistical identification with potentially endogenous proxies. arXiv:2302.13066 [econ.EM]
- Kilian L, Lütkepohl. 2017. *Structural Vector Autoregressive Analysis*. Cambridge, UK: Cambridge Univ. Press
- King M, Sentana E, Wadhvani S. 1994. Volatility and links between national stock markets. *Econometrica* 62(4):901–33
- Klein R, Vella F. 2010. Estimating a class of triangular simultaneous equations models without exclusion restrictions. *J. Econom.* 154(2):154–64
- Lanne M, Liu K, Luoto J. 2023a. Identifying structural vector autoregression via leptokurtic economic shocks. *J. Bus. Econ. Stat.* 41(4):1341–51
- Lanne M, Liu K, Luoto J. 2023b. *Identifying structural vector autoregressions via non-Gaussianity of potentially dependent structural shocks*. Work. Pap., Univ. Helsinki, Helsinki, Finl.
- Lanne M, Luoto J. 2020. Identification of economic shocks by inequality constraints in Bayesian structural vector autoregression. *Oxford Bull. Econ. Stat.* 82(2):425–52
- Lanne M, Luoto J. 2021. GMM estimation of non-Gaussian structural vector autoregression. *J. Bus. Econ. Stat.* 39(1):69–81
- Lanne M, Lütkepohl H. 2008. Identifying monetary policy shocks via changes in volatility. *J. Money Credit Bank.* 40(6):1131–49
- Lanne M, Lütkepohl H, Maciejowska K. 2010. Structural vector autoregressions with Markov switching. *J. Econ. Dyn. Control* 34(2):121–31
- Lanne M, Meitz M, Saikkonen P. 2017. Identification and estimation of non-Gaussian structural vector autoregressions. *J. Econom.* 196(2):288–304
- Lanne M, Saikkonen P. 2007. A multivariate generalized orthogonal factor GARCH model. *J. Bus. Econ. Stat.* 25(1):61–75
- Lee A, Hoesch L, Mesters G. 2022. *Locally robust inference for non-Gaussian SVAR models*. Work. Pap. 1357, Barcelona Sch. Econ., Barcelona, Spain
- Lee A, Mesters G. 2021. *Locally robust inference for non-Gaussian linear simultaneous equations models*. Work. Pap. 1278, Barcelona Sch. Econ., Barcelona, Spain
- Lewbel A. 1997. Constructing instruments for regressions with measurement error when no additional data are available, with an application to patents and R&D. *Econometrica* 65(5):1201–13
- Lewbel A. 2012. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *J. Bus. Econ. Stat.* 30(1):67–80
- Lewis DJ. 2021. Identifying shocks via time-varying volatility. *Rev. Econ. Stud.* 88(6):3086–124
- Lewis DJ. 2022. Robust inference in models identified via heteroskedasticity. *Rev. Econ. Stat.* 104(3):510–24
- Lütkepohl H, Meitz M, Netšunajev A, Saikkonen P. 2021. Testing identification via heteroskedasticity in structural vector autoregressive models. *Econom. J.* 24(1):1–22
- Lütkepohl H, Milunovich G. 2016. Testing for identification in SVAR-GARCH models. *J. Econ. Dyn. Control* 73:241–58
- Lütkepohl H, Milunovich G, Yang M. 2020. Inference in partially identified heteroskedastic simultaneous equations models. *J. Econom.* 218(2):317–45
- Lütkepohl H, Netšunajev A. 2015. *Structural vector autoregressions with heteroskedasticity: a comparison of different volatility models*. Discuss. Pap. 1464, Ger. Inst. Econ. Res., Berlin
- Lütkepohl H, Netšunajev A. 2017. Structural vector autoregressions with smooth transition in variances. *J. Econ. Dyn. Control* 84:43–57
- Lütkepohl H, Schlaak T. 2018. Choosing between different time-varying volatility models for structural vector autoregressive analysis. *Oxford Bull. Econ. Stat.* 80(4):715–35
- Lütkepohl H, Schlaak T. 2022. Heteroscedastic proxy vector autoregressions. *J. Bus. Econ. Stat.* 40(3):1268–81
- Lütkepohl H, Woźniak T. 2020. Bayesian inference for structural vector autoregressions identified by Markov-switching heteroskedasticity. *J. Econ. Dyn. Control* 113:103862
- Magnusson LM, Mavroeidis S. 2014. Identification using stability restrictions. *Econometrica* 82(5):1799–851
- Matilainen M, Nordhausen K, Oja H. 2015. New independent component analysis tools for time series. *Stat. Probab. Lett.* 105:80–87
- Matteson DS, Tsay RS. 2017. Independent component analysis via distance covariance. *J. Am. Stat. Assoc.* 112(518):623–37

- Maxand S. 2020. Identification of independent structural shocks in the presence of multiple Gaussian components. *Econom. Stat.* 16:55–68
- Mertens K, Ravn MO. 2013. The dynamic effects of personal and corporate income tax changes in the United States. *Am. Econ. Rev.* 103(4):1212–47
- Mesters G, Zwiernik P. 2024. Non-independent components analysis. *Ann. Stat.* 52(6):2506–28
- Milunovich G, Yang M. 2013. On identifying structural VAR models via ARCH effects. *J. Time Ser. Econom.* 5(2):117–31
- Moneta A, Entner D, Hoyer PO, Coad A. 2013. Causal inference by independent component analysis: theory and applications. *Oxford Bull. Econ. Stat.* 75(5):705–30
- Moneta A, Pallante G. 2022. Identification of structural VAR models via independent component analysis: a performance evaluation study. *J. Econ. Dyn. Control* 144:104530
- Montiel Olea JL, Pflueger C. 2013. A robust test for weak instruments. *J. Bus. Econ. Stat.* 31(3):358–69
- Montiel Olea JL, Plagborg-Møller M, Qian E. 2022. SVAR identification from higher moments: Has the simultaneous causality problem been solved? *AEA Pap. Proc.* 112:481–85
- Nakamura E, Steinsson J. 2018. High-frequency identification of monetary non-neutrality: the information effect. *Q. J. Econ.* 133(3):1283–330
- Normandin M, Phaneuf L. 2004. Monetary policy shocks: testing identification conditions under time-varying conditional volatility. *J. Monet. Econ.* 51(6):1217–43
- Pham DT, Garat P. 1997. Blind separation of mixture of independent sources through a quasi-maximum likelihood approach. *IEEE Trans. Signal Proc.* 45(7):1712–25
- Primiceri GE. 2005. Time varying structural vector autoregressions and monetary policy. *Rev. Econ. Stud.* 72(3):821–52
- Prono T. 2014. The role of conditional heteroskedasticity in identifying and estimating linear triangular systems, with applications to asset pricing models that include a mismeasured factor. *J. Appl. Econom.* 29(5):800–24
- Reiersøl O. 1950. Identifiability of a linear relation between variables which are subject to error. *Econometrica* 18(4):375–89
- Rigobon R. 2003. Identification through heteroskedasticity. *Rev. Econ. Stat.* 85(4):777–92
- Rigobon R, Sack B. 2003. Measuring the reaction of monetary policy to the stock market. *Q. J. Econ.* 118(2):639–69
- Rigobon R, Sack B. 2004. The impact of monetary policy on asset prices. *J. Monet. Econ.* 51(8):1553–75
- Sentana E. 1992. *Identification of multivariate conditionally heteroskedastic factor models*. Discuss. Pap. 139, Lond. Sch. Econ., London
- Sentana E, Fiorentini G. 2001. Identification, estimation and testing of conditionally heteroskedastic factor models. *J. Econom.* 102(2):143–64
- Shimizu S, Hoyer PO, Hyvärinen A, Kerminen A. 2006. A linear non-Gaussian acyclic model for causal discovery. *J. Mach. Learn. Res.* 7(72):2003–30
- Sims CA. 1980. Macroeconomics and reality. *Econometrica* 48(1):1–48
- Sims CA. 2002. Comment on Cogley and Sargent's evolving post-World War II U.S. inflation dynamics. *NBER Macroecon. Annu.* 16:373–79
- Skitovich V. 1953. On a property of the normal distribution. *Dokl. Akad. Nauk SSSR* 89:217–19
- Stock JH. 2008. *What's new in econometrics: time series*. Short Course Lecture 7 presented at 29th NBER Summer Institute, Cambridge, MA
- Stock JH, Wright JH. 2000. GMM with weak identification. *Econometrica* 68(5):1055–96
- Uhlig H. 2005. What are the effects of monetary policy on output? Results from an agnostic identification procedure. *J. Monet. Econ.* 52(2):381–419
- Wei T. 2015. A convergence and asymptotic analysis of the generalized symmetric FastICA algorithm. *IEEE Trans. Signal Proc.* 63(24):6445–58
- Wright PW. 1928. *The Tariff on Animal and Vegetable Oils: Investigations in International Commercial Policies*. New York: Macmillan